Advancing Thinking on Evidence for Biological Event Investigations

Kelsey Lane Warmbrod

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington
2025

Reading Committee:
John Scott Meschke, Chair
Gigi Gronvall
Deborah Bowen
Dongsheng Zang
Alex Greninger

Program Authorized to Offer Degree:
Public Health Genetics

University of Washington

Abstract

Advancing Thinking on Evidence for Biological Event Investigations

Kelsey Lane Warmbrod

Chair of the Supervisory Committee:

John Scott Meschke
Department of Environmental & Occupational Health Sciences

Biological events pose serious threats to public health, national security, and global stability. Identifying the origins of a biological event is critical for effective response and prevention of future events. However, biological attribution remains underdeveloped, lacking a conceptual framework and trusted investigatory processes internationally. This dissertation advances thinking about evidence for biological attribution by examining how diverse stakeholders perceive, interpret, and weigh evidence in the context of biological event investigations and by developing a conceptual framework to guide future efforts.

Using a mixed methods approach, this research draws on expert interviews in the scientific, public health, nonproliferation, and policy domains, as well as scenario-based exercises that tested how participants evaluated and prioritized different types of evidence. The findings reveal that across disciplines there are different views on the scope, feasibility, and the role of international organizations for biological attribution. Results highlight tension between scientific complexity and political realities, particularly concerning evidence collection, laboratory analysis, and communication of findings. Scenario exercises demonstrate that while genetic and laboratory data are often considered highly probative, other forms of evidence, such as epidemiological patterns or intelligence reports, are critical in shaping perceived credibility, **intent,** and actionability.

From these insights, a conceptual framework is proposed that integrates technical capabilities, investigation strategies, and contextual social and political factors. This framework underscores that attribution is

not only a scientific determination, but a multidisciplinary endeavor that requires the convergence of diverse evidence streams, transparent processes, and international confidence building.

By systematically exploring the evidentiary, social, and policy dimensions of attribution, this dissertation contributes to filling a critical gap in global biosecurity scholarship. The work offers practical guidance for designing future attribution investigations, informs ongoing debates about deterrence and accountability, and lays the groundwork for developing international norms and frameworks that can enhance preparedness, trust, and resilience in the face of biological threats.

Acknowledgements

First, I would like to thank all my committee members for their guidance and feedback through this process. This research would not have been possible without Dr. Gigi Gronvall and the late Dr. Deborah Bowen who helped tremendously in conceptualizing the project, helping find funding, and getting through IRB approval. Dr. Scott Meschke bravely took me on as an advisee and kept me on track through this process. Professor Donhsheng Zang has provided important insight on the legal insights affecting attribution. Dr. Alex Greninger has been particularly helpful in assessing technical applications for attribution. Finally, thank you to Robert Pekkanen for serving as the graduate school representative and teaching an amazing class on international relations.

Thank you so much to my former coworkers who helped make this project a reality. Nancy Connell provided invaluable guidance for this project and many others. Nancy and Gigi have been amazing mentors and friends throughout this process and have taught me so much about science, policy, and navigating the professional world. Another mentor has been Michael Montague, who dedicated hours to help think through problems and brainstorm creative solutions. Marc Trotochaud, Amanda Mui, Sanjana Ravi, Divya Hosangadi, and Christina Potter, have responded to every text, helped make figures or develop methodologies, and provided practical feedback as I was trying to get this project off the ground. Along with the others mentioned, Rachel West, Clint Haines, Natasha Kushnal, Lucia Mullen, Diane Meyer, Matt Shearer, Matt Watson, Crystal Watson, Rachel Vahey, and Tara Kirk Sell helped set up pilots of the survey and exercises and offered great feedback to improve the project. Mary Lancaster, Samuel Ortega, and Jonathan Forman were always available to chat about new ideas and research directions.

Thank you to the many participants who participated in one or more of the studies in this dissertation. The IRB does not allow me to name you, but your time and effort are greatly appreciated. Your dedication

to your fields and improving public health around the globe has been inspiring through this process.

Thank you to the many wonderful teachers and mentors I have had throughout my education. Mrs. Faye Curd first nurtured my interest in science back in 2nd grade. Mrs. Nayeema Ahmad encouraged my passion in high school. Dr. Cheryl Bast has mentored and supported me for over a decade now. These women, along with Drs. Connell and Gronvall, have shown me what it takes to be a woman in science and policy and to navigate this space with brilliance, humor, generosity, and grace. I would not be in a doctoral program without each of them. I can only hope to be half as good of a scientist, mentor, and teacher as they have been.

Thank you to all the friends who have helped me grow. There have been countless people along the way, most notably Emeline, Cassidy, David, Loni, Kate, Emily B, Paige, Caroline, Nathan, Tim, Emily K, Kathleen, Liz, Alex, Katie, Sarah, Stacy, Tracey- Ann, Rachel, Tiffany, Ian, Ashley, Elizabeth, Emily D, Adam, Luke, Elias, Angel, Matt H, Divya, Amanda, Sanjana, Lucia, Christina, Matt, Crystal, Abby, and Madison. Even if we have lost touch, they were a major part of my life, taught me many lessons, and were good friends when I needed them the most.

Thank you to my dogs, Teddy and Rufus, who have put up with my chaos while I have been working on this degree. They have dealt with five moves (two cross-country), five job changes, two pregnancies, two c-section recoveries, and two new children over the last 5 years, but never failed to be thrilled to see me when I walk in the door. They are the best company and endlessly entertaining.

Finally, thank you to my family, both found and by blood. I would not be who I am today without each one of you. You have been my rock throughout my life, but especially the last 5 years. Through the highs and lows, you have been my strength. From the bottom of my heart, thank you for being you and giving me the privilege of knowing and loving you: Kubra, Emily, Elizabeth, Taylor, Marc, Daniel, and Matthew. Thank you Nana for unconditionally loving me. Thank you Papa for teaching me courage and integrity and being my anchor. Thank you to Dad for always being patient and kind. Thank you to Mom for being my most steadfast support and teaching me strength and resilience. And countless hours of childcare to get this beast finished.

Thank you to my children Liliana and Mason. The greatest honor of my life is to be your mom. You are the greatest gift.

DEDICATION

To David Lane Mason Sr.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Biological risks, such as infectious diseases, pose a threat to public health, national security, and economies worldwide. The World Health Organization (WHO) reports over 7 million deaths from COVID-19 in 5 years. The International Monetary Fund (IMF) estimated there was a 5.2% shrinkage of global gross domestic product attributable to the COVID-19 pandemic (1). The risks posed by infectious agents must be minimized. As biotechnology advances and becomes more accessible, the risk of biological agents being misused also grows. Tools for minimizing the risk of a deliberate or accidental biological event, both technical and policy based, must continue to be developed as the risk potential grows.

Attribution refers to identifying the source of a biological agent, whether it be a laboratory the agent was manipulated in or actors that supported its development, storage, or dissemination. In chemical and nuclear nonproliferation, attribution is a deterrent for developing or using a weapon. Attribution in the biological context has not been well studied due to technical limitations, but recent advances in molecular biology and machine learning are being explored to advance attribution capabilities (2). However, there is no framework for using these new technologies in an investigation were a biological event to occur. There is similarly a lack of scholarship into how an investigation could be conducted to maximize trust in the process across disciplines and nationalities. While other categories of weapons of mass destruction have entire bodies dedicated to preventing risks associated with these materials through international treaties, there is no such international body for the implementation of the treaty preventing biological agent misuse. As such, attribution for biological agents has not been well studied and there is no clear implementation mechanism

for an investigation into the origins of an accidental or deliberate biological event, creating a dire gap in our ability to prevent a biological event with severe health, economic, and societal consequences.

Frameworks for forensic activities exist for cybercrime, nuclear activity, and chemical forensics (3–8). There are no conceptual frameworks specific to investigating the deliberate or accidental use of biological agents. This dissertation explores current views on the feasibility and potential of biological attribution, assesses how the evidence will be weighed by stakeholders, and understand how different factors could affect the credibility and willingness to act on results of an investigation. The results of this work will provide evidence and bounds for technologists and policy makers to guide their efforts to develop novel tools and policy initiatives for biological attribution investigations. The goal of this scholarship is to inform the development of a usable framework for conducting biological attribution investigations to better deter and prepare for biological events.

This dissertation uses a multi-method approach to investigate factors that will impact biological attribution including expectations of stakeholders, trustworthiness and reliability of evidence, and the relationships between different types of evidence. The following are the research questions that motivate this scholarship.

Research Question 1: Would an investigation into a biological event be understood, believed, and actionable?

Aim 1: Identify the attitudes of the international community concerning the current state of biological attribution and biological event investigations.

In Chapter 3, findings from key informant interviews with experts in various fields related to public health, microbiology, microbial forensics, international relations, and non-proliferations are presented. The goal of the interviews was to understand how biological attribution could help minimize risks of biological events, current capabilities for biological attribution, and expectations for how biological investigations should be done. The experts' expectations of biological attribution, evidence for attribution, and investigations are organized and categorized by stage of the investigation. Key pitfalls that could derail an investigation are identified. Implications for investigation design and best practices are included.

Research Question 2: Are different types of evidence generated in an investigation inherently weighted more heavily than others?

Aim 2: Assess beliefs and understanding of experts towards different evidence presented in fictious case

studies.

In Chapter 4, results from an exercise asking experts to judge evidence developed during an investigation are reported. Two fictious scenarios describing biological events were developed. Experts from different fields were asked to consider evidence found in investigating these events and respond to a survey assessing how they used different types of evidence when drawing a conclusion about the biological events. Participants were asked to rank the evidence. Through this exercise, we gain an understanding of how different types of evidence are judged and weighted. Considerations that could impact the trustworthiness, weight, and actionability derived from different types of evidence are highlighted in this chapter.

Research Question 3: What factors influence the actionability and acceptability of an investigation?

Aim 3: Examine the relationship between the social, technical, and legal aspects of biological investigations.

Chapter 5 considers the relationships between factors influencing biological investigations and the evidence they generate. Results from the interviews and scenarios informed a conceptual framework illustrating the relationship between technical, social, and political factors influencing perceptions and trust of evidence and investigations for attribution. This information can inform the development of protocols for investigations and guide investigatory teams in how to proceed through an investigation.

Chapter 2

Literature Review

2.1 Biological Events

When a disease spreads among human, animal, or plant populations, there can be significant impacts on population health. Communicable diseases caused by microbial organisms can enter a population through natural transmission pathways or anthropogenic means, such as an accidental or deliberate event. Understanding how the biological agent entered the population can be important in preventing future events and implementing interventions to mitigate further harm (9).

Regardless of how a biological event begins, the role of public health in response is critical. A key goal will be to find out how the disease is entering the population. This is important for several reasons, such as control and prevention of further transmission, information about treatment options, and preparedness for future events (10).

Understanding how a disease emerges and spreads within a population helps us understand transmission pathways and spillover risks, enabling better risk assessments to prevent future events (11). Origins of a pathogen can also inform treatment options (12). Part of understanding where a pathogen came from is understanding the evolutionary history of the agent. This information can help determine whether there are existing medications that could be effective against the new agent on the basis of its similarity to a prior agent.

2.1.1 Naturally Occurring Events

There are different characteristics that help determine the origin of an event. For example, how it was detected may indicate whether it was natural, accidental, or deliberate (13).

Naturally occurring outbreaks arise from the interaction between pathogens and their natural hosts within ecosystems. These outbreaks can be caused by viruses, bacteria, fungi, or parasites, and they often emerge due to ecological changes, population dynamics, or genetic mutations within the pathogen. Outbreaks arising from consuming contaminated food or water, bites from infected animals or vectors, or person-to-person transmission are all naturally occurring outbreaks.

Notable naturally occurring outbreaks include the Black Death in Europe and Asia from 1347-1351, the 1854 cholera outbreak in London, the 1918 Influenza Pandemic, the HIV/AIDS pandemic starting in the 1980s, and the West African Ebola outbreak from 2014-2016. For each of these outbreaks, there were challenges with identifying the origin of the outbreak.

The identity of the agent causing the Black Death was not known in the 14^{th} century. *Yersinia pestis* wasn't even discovered until 1894 and it was only in 2010 that *Y. pestis* DNA was isolated from bones of individuals who died during the Black Death to definitively confirm that it was the agent causing the pandemic (14; 15). Only in 2022 did researchers publish their determination that the strain of *Y. pestis* that caused the Black Death evolved in what is today Kyrgyzstan (16). During the pandemic, people had many ideas about where the disease had come from including divine punishment, astrological origins, and fumes from bad air, known as the miasma theory (17). Lack of technology and knowledge of microbiology and epidemiology were major barriers to identifying the origin of the Black Death at the time of the pandemic.

During the 1854 cholera outbreak in London, the initial two theories of where the sickness was coming from was the miasma theory and what would later be known as germ theory. Physician John Snow famously mapped the homes of people who died from the disease on a map and traced it back to a water pump, concluding that the disease was coming from something in the water from that pump and founding the field of epidemiology (18). Meanwhile, the bacterium *Vibrio cholerae* was isolated for the first time the same year, but not widely accepted as the cause of cholera for many years after the outbreak (19).

The origins of the 1918 influenza pandemic continue to be a point of debate. The first recognized outbreaks of this pandemic occurred in March 1918 in US military camps. However, there are some who

argue that outbreaks occurred earlier in other regions of the world, including France or China (20). Prior to March 1918, there had been a severe respiratory illness circulating among workers on the Western Front of World War I. At the time of the pandemic, there were no viral isolates taken and studied, which has complicated efforts to identify the pandemic's origin. Researchers today rely on historical records and genetic reconstruction of the virus from preserved tissues (21; 22). The pandemic was named the "Spanish flu" because Spain was neutral in the war, which allowed more free reporting about the pandemic in countries participating in the war, illustrating how geopolitics can impact response to public health emergencies (23).

In modern times, epidemiologists have looked for a patient zero to determine the "origin" of an outbreak (24–26). By identifying the first infected patient, it can be easier to trace where the initial person became infected. However, there are ethical questions surrounding how this individual is identified and discussed (27). There are also challenges with identifying a true patient zero when modeling outbreaks that were not immediately detected (28). Even if a patient zero is identified, finding the source of their infection can be challenging. Travel history, contact tracing, and ingestion logs can help narrow down potential sources of infection for the individual. If a source of infection for patient zero can be identified, there may still be a question of how the pathogen evolved or emerged (29). Did it evolve in an animal and spillover into the human population? If so, what is the reservoir animal? If there is no direct contact between the reservoir and humans, then how did it jump from the animal population to the human population? Is there an intermediate host and, if so, what is it and what is the transmission cycle between all three animals (30)? It can be very complex to determine an origin.

Additionally, as technical capabilities expand or more data is gathered, attribution theories may change. In the early 2000s, there was debate among scientists over the origins of the pathogen responsible for most severe form of malaria, *Plasmodium falciparum*. One team suggested it has an ancient human origin after isolating ancient DNA of the parasite from Egyptian mummies (31). Another team suggested the pathogen had evolved in gorillas (32). Since then, scientists have come to agree on a gorilla origin based on more sampling and analysis of samples from gorillas (33). While this is a normal part of the scientific process, attribution for biological events will not just be a scientific endeavor. Public health, which touches the lives of most people much closer than general research, is also impacted by attribution. Evolving attribution theories can therefore be vulnerabilities for public

health. In high pressure situations, like health threats, a changing theory of where the pathogen came from can be a prime target for bad actors to exploit fears or uncertainty. Mis- and dis-information can thrive in these environments.

2.1.2 Accidental Events

Accidental outbreaks are typically the result of human activities or errors that lead to the release of harmful agents beyond the laboratory. Accidental outbreaks can arise from laboratory-acquired infections (LAIs) and accidental pathogen escape from laboratory settings (APELS) from breaches in containment protocols, unsafe handling of hazardous materials, or technological failures (34). Accidental outbreaks could also occur due to a natural disaster causing a failure of containment or from unintended consequences of benign use of a pathogen. A 2024 literature review found reports of 309 LAIs and 16 APELS (35). The APELS were primarily attributed to procedural errors such as using expired disinfectants when manufacturing vaccines, which would fall under the category of an unintended consequence of a benign use of a pathogen.

One key distinction between naturally occurring and accidental outbreaks lies in their predictability and controllability. Naturally occurring outbreaks, while unpredictable in terms of timing and magnitude, are often influenced by identifiable ecological, epidemiological, and climatic factors. Surveillance systems, early warning mechanisms, and public health interventions can help detect and mitigate these outbreaks, although challenges may arise from the complexity of natural ecosystems. In contrast, accidental outbreaks are often preventable and attributable to specific human actions or lapses in safety protocols (36). These outbreaks highlight the importance of risk assessment, regulatory oversight, and adherence to best practices in various domains such as laboratory safety, industrial hygiene, and transportation security. Effective prevention and response to accidental outbreaks require robust emergency preparedness measures, rapid detection capabilities, and coordinated response efforts across multiple stakeholders.

2.1.3 Deliberate Events

A deliberate event, often referred to as a bioterrorist or intentional act, involves the deliberate release or use of biological agents, such as bacteria, viruses, or toxins, with the intent to cause harm, instill fear, or disrupt societal functions. These acts may be perpetrated by individuals, groups, or state actors seeking

to achieve political, ideological, or strategic objectives. A deliberate event could include an assassination, terrorist attack, warfare, intentional experimentation on an individual or population with a biological agent, or another criminal act.

One of the key distinctions between deliberate and accidental outbreaks lies in their underlying motives and intent. Deliberate outbreaks are characterized by malicious intent and premeditated actions with the goal of causing harm. Objectives of a deliberate event could include causing terror and chaos, undermining public confidence, or exerting political pressure. In contrast, accidental outbreaks result from unintended consequences or unforeseen circumstances, often stemming from human fallibility, system vulnerabilities, or natural phenomena. Furthermore, deliberate outbreaks are typically orchestrated with secrecy, planning, and coordination to maximize their impact and evade detection, whereas accidental outbreaks occur spontaneously or because of isolated incidents, lacking the intentional design and coordination associated with deliberate acts (37).

Examples of deliberate events include the 2001 Amerithrax letters, intentional food poisoning of salad bars to influence local elections by the Rajneeshee community in Oregon in 1984, and actions of so-called Unit 731 of the Japanese military during World War II. In the 2001 Amerithrax letters case, it is believed that scientist working for a government laboratory at Fort Detrick, Maryland stole anthrax from the laboratory and sent it through the mail to senators and members of the media (38). Seventeen people were exposed and five died in connection to the letters. The FBI lead a several years long investigation into the letters, followed by a National Academies of Sciences report that questioned the science justifying the findings of the FBI, to which the FBI responded, but to date some people still do not consider the matter settled.

The Rajneeshee cult in The Dalles, Oregon wanted to gain political control over their local circuit county court so before the local elections in 1984, they put salmonella on ten local salad bars across the county. There was a total of 751 people who got salmonellosis in connection with the cult's actions. Initially, it was believed that the outbreak was caused by poor hygiene from food handlers. However, one congressman continued pressing and argued three circumstantial points on the floor of the US House of Representatives to keep an investigation going. He argued that the abnormally large size of the outbreak, the presentation of cases in two waves, and the fact that there was more than one source where people were infected suggested there was an intentional aspect to the outbreak (39). Months later he was proven right by the investigation.

Unit 731 was a Japanese military unit in World War II that conducted experiments on primarily Chinese and Russian citizens that were kidnapped or prisoners of war (40). Among the many types of experiments done on the kidnapped were injections of pathogens including syphilis, gonorrhea, smallpox, *V. cholerae*, and others. Following infection, the effects of the infections were studied and, in some cases, published in peer-reviewed literature (41). The experiments done by Unit 731 were used to develop biological weapons for the Japanese military to use in China during World War II.

2.2 Biological Weapons

Biological weapons have two components, an agent and a delivery mechanism. Unlike conventional weapons, biological agents can be administered through various means, including ingestion, inhalation, or contact with contaminated surfaces, allowing perpetrators to evade detection and attribution (42).

Biological weapons can be used as instruments of terrorism to instill fear, panic, and social disruption by causing mass casualties, illness, or economic damage. Terrorist groups or individuals may seek to exploit the accessibility and lethality of biological agents to achieve ideological, political, or religious objectives. Biological terrorism poses unique challenges due to the potential for indiscriminate harm, psychological impact, and the difficulty of detecting and preventing covert attacks. Examples of biological terrorism include the 1995 Aum Shinrikyo cult's attempt to release anthrax in Tokyo and the 2001 anthrax attacks in the United States (38; 43; 44). Toxins like ricin and botulinum have been used to assassinate political dissenters (45; 46).

Like their use in terrorism, biological weapons can similarly serve as tools of psychological warfare to demoralize, intimidate, or coerce adversaries by exploiting fear, uncertainty, and vulnerability. The mere threat of a biological attack or the perception of biological hazards can have profound psychological effects on individuals, communities, and societies, undermining confidence in government institutions, eroding social cohesion, and exacerbating public anxiety (47). Psychological warfare tactics involving biological threats may include propaganda, disinformation campaigns, or simulated bioattacks designed to manipulate perceptions and behavior.

Biological weapons have been employed as instruments of warfare several times, often in conjunction with other forms of military force to achieve strategic or tactical objectives. Biological warfare involves

the deliberate use of biological agents to incapacitate or kill enemy combatants, disrupt military operations, or undermine civilian morale and infrastructure. Biological weapons offer several advantages in warfare, including their potential for mass casualties, low cost of production, and covert delivery methods (48). Historical examples of biological warfare include the use of plague-infected corpses during sieges and conflicts in ancient times (49).

Biological weapons can also be used for criminal purposes, such as extortion, sabotage, or financial gain. Biocrime involves the illicit acquisition, production, or dissemination of biological agents to perpetrate criminal acts, such as targeted attacks on individuals or businesses, sabotage, or bioterrorism for personal or organizational motives (50). Biocrime poses challenges for law enforcement and public health authorities due to the clandestine nature of these biological threats and the potential for criminal organizations or individuals to exploit vulnerabilities in biosecurity and biosafety protocols (51).

Biological weapons are inherently unique compared to the other non-conventional weapons. Unlike chemical and nuclear materials, many biological agents are infectious or capable of spreading beyond the control of humans, like gene drives. Biological agents are also capable of evolution, which complicates our understanding of how such agents function and change, the potential for detecting and tracking of agents, and applicability of countermeasures (52). Creating a biological weapon also doesn't necessarily require a facility that satellite images can monitor (53). Because of these characteristics, it is hard to detect a biological weapons program or respond to the use of a biological weapon.

Also challenging for the biological weapons space is determining what is or is not a biological weapon. Biological sciences and technology are inextricably linked to public health and medicine, much more than nuclear and chemical sciences. There are also other peaceful purposes for biology like environmental remediation, agriculture, and the bioeconomy. The links to health, the environment, and food means there will always be significant peaceful purposes for pursuing advancements in this field. Life sciences touch all aspects of day-to-day life, generating many opportunities for economic development and solving challenges via biology. With all these peaceful uses of biology, the existence of a biological agent or the research activities themselves doesn't necessarily mean a biological weapon exists or is being created (54). Determining biological weapons status therefore becomes a question of intent, rather than of activity or product, which can be much more challenging to determine (55).

2.3 Nonproliferation

The proliferation of WMD poses a significant threat to international security and stability. Acquisition, production, stockpiling, or use of such weapons (including nuclear, chemical, and biological weapons) by state and non-state actors continue to be challenges for the international community. There are fundamental differences between each of the types of WMD. Biological weapons have the potential to spread widely between people after initial release (56). Nuclear weapons can have devastating effects for decades following use, but it is generally limited to the area surrounding deployment of the weapon over the long term. Chemical weapons can similarly cause massive damage acutely but typically do not travel from person to person. Weather patterns and payload sizes can affect how far ranging and long lasting nuclear and chemical weapons can be dangerous, but neither has the potential for sustained, long term continuous transmission unlike some biological agents (57).

There are several international treaties dedicated to nonproliferation of WMD. In 1928, the Geneva Protocol (formerly known as the Geneva Protocol for the Prohibition of the Use in War of Asphyxiating, Poisonous or Other Gases, and of Bacteriological Methods of Warfare) entered into force and banned the use of chemical and biological warfare. Notably, this convention banned the use but not research nor production of such weapons and allowed for the use of such weapons against non-signatory states (58). The Biological and Toxins Weapons Convention (BWC) from 1972 banned states from developing, producing, acquiring, or stockpiling biological or toxin weapons entered into force in 1975 (59).

International treaties govern nuclear and chemical nonproliferation as well as biological nonproliferation. The Nuclear Non-proliferation Treaty (NPT) opened for signature in 1968, entered into force in 1970, and currently has 191 signatories. There are three main objectives of the treaty: prevent the spread of nuclear weapons and other nuclear explosive devices by preventing states that don't currently have nuclear weapons from pursuing them, disarmament of nuclear weapons in states with such weapons, and affirming the right of all signatories to use nuclear technology for peaceful purposes like energy production (60). The International Atomic Energy Agency (IAEA), an independent international organization that works in cooperation with the UN to promote the peaceful use of nuclear energy while preventing the proliferation of nuclear weapons, monitors compliance with nuclear non-proliferation treaties such as the NPT. IAEA also implements safeguards agreements with member states to verify the peaceful use of nuclear materials and

facilities, conducts inspections, and provides technical assistance to strengthen nuclear security and safety standards (61).

The Chemical Weapons Convention (CWC) opened for signatures in 1993 and entered into force in 1997. The CWC includes provisions for eliminating the use of chemical weapons, destroying existing stockpiles of such weapons, and promoting peaceful use of chemistry in areas such as medicine or agriculture. The Organization for the Prohibition of Chemical Weapons (OPCW) was created by the treaty and verifies compliance with the CWC through inspections, monitoring, and verification activities, and it assists member states in destroying chemical weapons stockpiles and related facilities. The OPCW also promotes international cooperation and assistance to strengthen chemical safety and security measures (62). Both the NPT and CWC apply to nation states. In 2004, the United Nations Security Council Resolution 1540 (UNSCR 1540) was agreed to with the aim of prohibiting the use of nuclear, chemical, and biological weapons by non-state actors (63).

The United Nations Secretary-General's Mechanism for Investigation of Alleged Use of Chemical and Biological Weapons is a vital component in the international framework for nonproliferation and disarmament efforts. Established in 1987, this mechanism provides for the prompt investigation of allegations of the use of chemical or biological weapons at the discretion of the UN Secretary General. Once triggered, the mechanism allows for a thorough investigation into allegations of use of chemical or biological weapons (64). The United Nations Office for Disarmament Affairs (UNODA) serves as the focal point within the UN Secretariat for disarmament matters, including nuclear, chemical, and biological disarmament, including activation of the UNSGM. UNODA promotes multilateral negotiations, facilitates dialogue among member states, and provides substantive support for disarmament efforts, including the implementation of relevant treaties and agreements.

2.3.1 Verification

Verification is a key piece of nonproliferation strategy. Verification refers to processes designed to confirm compliance with relevant treaties and norms. Verification may include inspections, monitoring, surveillance, or exchange of scientific research and best practices (65). When successful, verification is meant to increase transparency amongst states, increase confidence in safety and security measures, and improve trust between

states and the nonproliferation regime. In the nuclear and chemical disarmament area, there have been many programs dedicated to verification that nuclear or chemical materials have not been misused.

On the nuclear side, to meet NPT requirements, there are IAEA inspectors who will inspect facilities and verify there are not weapons grade materials in a facility. Non-nuclear states submit a comprehensive safeguards agreement with the IAEA in which they declare all nuclear material and facilities in their country (66). The IAEA then conducts inspections and investigates to determine if nuclear material is being used for weapons. Inspectors will use on-site inspections, environmental sampling, satellite imagery, records review for material accountancy, and tamper-proof seals to determine if a country is in compliance with the NPT. For the five recognized nuclear armed states under the NPT (USA, UK, France, Russia, and China), the verification methods are more focused on transparency and confidence building. Military sites do not have be to be inspected, but civilian facilities can inter voluntary offer agreements to support transparency (67). Nuclear armed nations have additional arms control agreements, mostly bilaterally, that may have additional verification requirements (68; 69).

The CWC similarly has inspectors through OPCW that can look for evidence of weapon production. Countries are supposed to declare what chemicals are used, their purpose, and quantity to OPCW (70). OPCW then sends inspectors to verify the information in the declarations. For both nuclear and chemical weapons use, the inspectors from IAEA and OPCW, respectively, can be tasked with investigating a chemical weapons use event if needed (71).

Unlike with chemical and nuclear verification, there is no formal verification mandate in the BWC (72). State parties are encouraged but not required to submit annual reports on activities that could be relevant to the BWC, like national legislation on biosafety or the presence of high containment laboratories. There was an effort to add a verification protocol to the BWC, known as VEREX, but that effort fell apart in 2001 when there was concern that inspections could threaten national security. There are efforts to increase trust in other ways, such as confidence-building measures (73).

2.3.2 Deterrence

Within the WMD community, significant work is done to advance deterrence of the production or use of a weapon. Deterrence is a strategy to dissuade actors from using a WMD, typically by making the cost of us-

ing such a weapon outweighing potential benefits (74). Deterrence may make choosing a more conventional weapon that has a lower potential for destruction more appealing to a bad actor by making the consequences if caught using a WMD much greater than a traditional weapon. Deterrence therefore supports nonproliferation efforts by changing the behavior of bad actors (75).

In nuclear security, attribution is considered a deterrent to using nuclear weapons. If actors considering using a nuclear weapon know there is a good chance authorities will be able to trace the weapon back to them, then they are less likely to use that weapon (76). The same theory could be applied to biological weapons, though it is unclear how strong a deterrent attribution is; there remains a potential that the adversary won't care if the action is attributed to them. However, even if the perpetrators of an attack don't care if the attack can be traced to them, attribution can still be beneficial to others (77). Just as there is a potential to determine who is responsible, there is also a potential to demonstrate someone else is not responsible (78). Attribution could result in evidence demonstrating an accused entity is not at fault. In this way, attribution can be used to build trust, especially between countries on the international stage. If attribution does conclusively identify the responsible party, the evidence gathered in the attribution investigation can be used for criminal proceedings to hold the responsible party accountable, though there is risk of a sophisticated actor spoofing attribution by designing an agent to look like something another entity made. Looking for a convergence of evidence from many different sources will therefore be important. Finally, results from attribution can be used to inform interventions to mitigate the potential for repeat events.

2.4 Defining Attribution

Despite growing interest in attribution in different venues, there is not a standard definition for the term or common agreement on the breadth of activity included in attribution. In public health, source attribution often refers to identifying the specific source of an infectious disease outbreak, most often with foodborne illnesses (79). Within the biosecurity literature, definitions have varied from "being able to determine, in the wake of a human-caused biological event, who was responsible," which includes the human-caused qualifier, to "determining who is responsible and culpable," which has no qualifier about the type of event (13; 80). The 2022 Biodefense Strategy mentions "timely and effective biological material characterization to support investigations, origin determination, and attribution," implying attribution is different from either

an investigation or origin determination (81).

Attribution could cover all the following questions: Where did the agent come from? Who created and/or manipulated the agent? Who funded the creation, storage, or dissemination of the agent? Who released the agent? Was the agent released on purpose? If all these questions are included in attribution, then any type of event should be covered (natural, accidental, deliberate). There is also the question of what activities may be included in attribution. Does attribution include initiating an investigation, sample collection, analysis, and reporting? Microbial forensics, which is often considered to be a discipline related to analyzing microbial evidence for attribution, is not a stand in term for attribution but could be considered part of attribution (82). If attribution was limited to microbial forensics, then this would exclude other types of evidence that could help determine where a biothreat came from such as intelligence signals or public health data.

Clarifying what is or is not included in attribution is important for strengthening the capability; funding decisions, appropriate stakeholder inclusion, and authority to conduct attribution are all affected by defining what is or is not attribution. For example, if attribution includes identifying a point source regardless of the type of event, then public health should be included in attribution. If attribution includes sample collection and analysis, then laboratories with expertise in different types of biological agents need to be included in planning for attribution. If entities are not aligned on their understanding of the term, there could be turf wars between agencies about who is responsible for attribution.

By clarifying what counts as attribution, frameworks for advancing the capability can begin to form. Without a shared understanding, it will be challenging to create new legal measures or policy frameworks to enable this work to be done. Defining what is included or not in attribution helps determine who should be responsible for leading research into advancing the capability and developing improved frameworks for implementing attribution.

The extent to which attribution can be effective on the international scale may be impacted by its definition. If attribution is narrowly defined as only applying to deliberate events and only determining who is responsible, it is likely to be considered something primarily done by law enforcement and intelligence entities (83). In such a case, countries may be less inclined to share their information or results from attribution with other countries, which could diminish the possible trust that could be built between countries. If attribution is considered a national security issue more than a public health issue, then public health agen-

cies that house critical data for determining source may be reluctant to participate as participation in law enforcement or intelligence activities could erode public trust in public health. For WHO, whose funding is highly dependent on voluntary contributions of different states, if one state finds WHO's activities to be inappropriate, they risk losing critical funding (84; 85).

The word attribution could be used to refer to the activities related to identifying a source and responsible parties following detection up until reporting. Acting upon results of attribution should be considered a different step because decisions concerning appropriate action will be based on much more than the results from attribution, such as politics and morals, and involve a much broader set of stakeholders, such as voters. Detection should also be a different step because detection will rely on a different set of tools. However, data from detection will be helpful in informing attribution, especially as we move towards more threat agnostic methods of detection (that instead of relying on the identity of the threat, rely on characterizing the host response**, which could lead to faster treatment of patients in a biological event**). Microbial forensics should be considered a tool for attribution and one piece of the puzzle. Intelligence signals, traditional criminal investigatory techniques, public health and medical records, and comparisons to 'normal' behavior for laboratories should all be considered as evidence to inform attribution in addition to microbial forensics. Looking for a convergence of results from each of these types of evidence will lower the risk of an adversary successfully "tricking" authorities.

To strengthen our national and international capability for attribution, we need to fund more research into microbial forensics and agent characterization, public health surveillance **(including the networks of public health, clinical, and reference laboratories),** and understanding the 'normal' behavior of laboratories. If this is to be used as a deterrent, we need to do better at talking about it so potential adversaries know we have the capabilities. We need to create frameworks that clearly define what is attribution and who is responsible for doing it and when, as well as identifying other entities that will be involved. **There must be strategies to handle data generated that cover data security, sharing, and publishing.** For the USG, there needs to be money given to agencies to research this area and organize the agencies around a common understanding of what attribution is.

The 2022 National Biodefense Strategy and Implementation Plan and 2022 National Security Strategy, both from the Biden- Harris Administration, both mention attribution of biological agents as goals for biode-

fense (81). Attribution is an important consideration for biodefense: attribution can be a deterrent from using a biological weapon, it can be used to build trust amongst the international community, through attribution we can learn new measures to prevent future events from happening, and results from attribution may be used for accountability. Despite the potential benefits of attribution, there are substantial gaps at both the international and national level related to implementing attribution (38; 82).

2.5 COVID- 19 Origins

Determining the origins of COVID-19 is a recent example of why attribution is important and the consequences of ambiguity in determining origin. Despite the pandemic beginning over five years ago, there continues to be contentious debates about the origins of the pandemic (86).

One of the first scientific publications to address possible origins of the SARS-CoV-2 virus was the "Proximal Origins" paper released in early 2020 (87). Here, viral genomes from early patients were compared to other coronaviruses. The authors concluded that the virus was unlikely to have been engineered and that features of the virus were consistent with naturally occurring evolution. Another 2020 paper identified high similarity between SARS-CoV-2 and a coronavirus previously isolated from bats, RaTG13, suggesting there could be a bat reservoir for SARS-CoV-2 (88). Both these papers support a natural occurrence hypothesis for the pandemic. Importantly, these papers were authored by experts in viral evolution, virology, molecular genetics, and microbiology. Subsequent publications have further studied SARS-CoV-2 evolution and found typical evolutionary patterns (89–92).

In response to the quick proliferation of conspiracy theories about COVID-19 origins, scientists who were experts in viral genetics and evolution wrote why naturally occurring origin was the most likely. In 2021, The joint WHO-China investigation released a report that assessed zoonotic spillover from an intermediate host to be "likely to very likely" and considered a laboratory origin to be "extremely unlikely" (93). However, this report was criticized for not giving adequate consideration to a possible lab leak and being heavily influenced by the Chinese government, who have a vested interest in not being blamed for causing a pandemic that killed millions of people (94).

There were additional publications that raised questions about transparency concerning data relevant to COVID-19 origins. One such paper revealed that early SARS-CoV-2 sequences had been deleted from the

NIH's Sequence Read Archive (95). Other papers were published that stressed that both the natural and laboratory origins should remain hypothesizes worthy of study (96).

Outside of the scientific literature, there was heated debate on social media that the origin could be an accidental laboratory escape or more nefarious, a bioweapons experiment gone wrong. Even the President of the United States suggested the virus was "made" in China in early 2020, without providing any evidence (97). These theories were repeated by politicians, conspiracy theorists, and some scientists, often tied to xenophobic remarks about China (98). Among scientists who publicly suggested COVID-19 was an accidental or deliberate event, most were not experts in virology, viral genetics, or viral evolution. One loud voice was a scientist who specialized in gene therapy and cell engineering, which does have some similar overlap to virology, but is ultimately not the same (99). Viruses, which must infect other cells to replicate, evolve very differently than multicellular organisms or even unicellular organisms (100–103). Another loud proponent of the lab leak theory was a molecular biologist who initially thought the lab leak theory couldn't be ruled out, but then went on to say a natural origin was impossible and called people who published evidence of a natural origin "provably coauthors of fraudsters and perjurers (104)."

Intelligence agencies also did analyses of the origins of COVID-19 (105). US intelligence assessments released under the COVID-19 Origin Act of 2023 showed a divided view amongst the intelligence community. The Department of Energy and the FBI considered a lab-related accident plausible, while other agencies leaned toward natural spillover, though all with low or moderate confidence (106). The report, released by the Subcommittee on the Coronavirus Pandemic, concluded that COVID-19 "likely originated" from a lab leak in Wuhan. However, in 2025, the CIA noted they had "low confidence" in their assessment that the lab leak occurred in Wuhan and added that China continues to impede origin investigations (107).

The COVID-19 origins debate was impacted by the long-standing debate about dual use research and gain of function (GoF) research. For years, there had been debate in the US concerning if GoF research should be allowed and what safeguards should be in place for conducting such research (108; 109). Because one of the primary theories for COVID-19 origins was a lab release, the debate on COVID-19 origins was often conflated with the GoF debate (110). In September 2025, Donald Trump said "Just a few years ago, reckless experiments overseas gave us a devastating global pandemic, yet despite that worldwide catastrophe, many countries are continuing extremely risky research into bio-weapons and man-made pathogens.

This is unbelievably dangerous (111)." People who strongly condemned GoF research, without having a reliable definition for what was or was not GoF, tended to become sympathizers of the lab leak theory (112–114).

In 2025, the White House released a "report" blaming the WHO, the Department of Health and Human Services, and former leaders of science policy, among others for "failures" during the pandemic (115). The report has no evidence included to back it up and, in several sections, includes statements that are completely contradicted by most research. The evidence cited in the report to support the assertion that the virus came from a lab in Wuhan are as follows (copied verbatim, including emphases):

1. The virus possesses a biological characteristic that is *not found in nature*.

2. Data shows that all COVID-19 cases stem from *a single introduction into humans*. This runs contrary to previous pandemics where there were multiple spillover events.

3. *Wuhan is home to China's foremost SARS research lab*, which has a history of conducting gain-of-function research (gene altering and organism supercharging) at inadequate biosafety levels.

4. Wuhan Institute of Virology (WIV) researchers were *sick with COVID-like symptoms* in the fall of 2019, months before COVID-19 was discovered at the wet market.

5. By nearly all measures of science, if there was evidence of a natural origin it would have *already surfaced*. But it hasn't.

The report does not say what biological characteristic is not found in nature. There are no peer reviewed citations that suggest there is something in SARS-CoV-2 that is not found elsewhere in nature.

First, scientists believe there were 2 spillover events for SARS-CoV-2, so statement 2 is false (116). Furthermore, there can be multiple spillover events to seed a pandemic or epidemic, but there are also several cases where scientists believe it was a single spillover that started the pandemic. In the case of HIV, it is believed that different strains of simian immunodeficiency viruses (SIVs) were spread from primates to humans on several occasions, but transmission chains beyond the initial infections were limited. The AIDS pandemic is believed to have started with one transmission of a SIV to humans, which led to the HIV-1 group M virus that could sustain transmission in humans, which caused the pandemic (117). Somewhat similarly, both the original SARS pandemic and MERS outbreaks post 2012 are believed to have occurred from repeated zoonotic introductions, but with most introductions leading to limited transmission. This means that

only a fraction of spillover events for MERS and the SARS-CoV are leading to sustained human-to-human transmission chains (118). The 1918 influenza pandemic may have been started from one spillover event or several, scientists aren't sure, but either is plausible (119). The 2009 H1N1 pandemic was believed to start from a single spillover event because many of the early genomes sampled from patients were homogenous (120). Each cross -species transmission event is a bottleneck event for a virus. A new host species means new tissues with different pressures the virions must contend with to seed an infection, creating a bottleneck for a virus population (121). A viral population that has undergone a bottleneck can have a lowered fitness level than the original population, which can impact its ability to create a sustained transmission chain (122).

GoF research is not a well-defined term nor concept. There are thousands of experiments done every year that alter a gene, some of which is done on purpose, and some is just the nature of working with a microbe in a laboratory. Intentional changes are often done using site directed mutagenesis for research, therapeutic, or commercial purposes (123–126). Simply passaging a virus can lead to random alterations in genes (127). This is the nature of biology. Gene altering, even when done intentionally, does not equate to nefarious intent. Additionally, "supercharging" is not a scientific term.

COVID-19 symptoms are similar to many respiratory tract infections, such as influenza, RSV, or the common cold in adults (128). The symptoms of COVID-19 are not unique to the SARS-CoV-2. The presence of symptoms that are similar across several diseases is not strong evidence for concluding a laboratory origin. Additionally, even if these laboratory workers were infected with COVID-19, that does not mean it came from the laboratory. They could have been infected in their community. Without looking at broader epidemiological data from Wuhan in the fall of 2019, it would be premature to conclude that these infections were from the laboratory.

As detailed above, there are many papers that detail evidence of a natural origin. Cherry-picking data or completely ignoring studies does not equate to there being no evidence.

Also in 2025, the WHO's Scientific Advisory Group for the Origins of Novel Pathogens (SAGO) released their final report on COVID-19 origins. The Interim Report from 2022 stressed that both hypotheses remained open, while the final 2025 report concluded that the "weight of available evidence" favors zoonotic spillover, either directly from bats or via an intermediate host, but also emphasized that critical gaps remain (129; 130). The absence of definitive evidence for an intermediate animal host, the lack of unrestricted

access to early case data, and unresolved biosafety questions at laboratories in Wuhan continue to sustain alternative interpretations for the origins.

2.6 Mis- and Disinformation

Misinformation, or information that is false or misleading and spread unwittingly by people unaware it is false, and disinformation, or information that is false and intentionally spread knowing it is false, have both impacted responses to health threats. During the COVID-19 pandemic, the WHO declared an "infodemic" based on the amount of mis- and disinformation being spread about COVID-19 that was severely impacting response efforts (131). Misinformation has similiarly harmed outbreak responses for Ebola (132), Zika (133), and malaria (134). Both forms of incorrect information can impair response efforts by eroding trust in public health efforts leading to decreased willingness to follow guidelines. Outbreaks, especially pandemics, are particularly vulnerable to misinformation due to enhanced uncertainty and inherit elevated sensitivities when discussing health.

Trust in science and public health is vital to outbreak response. Mis- and disinformation undermine trust by exploiting gaps in scientific knowledge (such as origins of a pandemic pathogen), amplifying uncertainties and presenting them as a failure of science or public health, and presenting false information from people who are not experts. Risk communication is already a vulnerability of disaster response because effective communication is highly dependent on context and audience. As events grow in size and fear increases, vulnerabilities also grow (135). This is most evident on social media websites. Social media has enabled unprecedented speed of information dissemination. Algorithms amplify emotionally charged content, such as information about pandemics, rapidly and with little moderation (136). Conspiracy theories can proliferate, especially when bots are utilized to manipulate algorithms to artificially amplify the message. Usage of bots during COVID-19 has been correlated with willingness to ignore mask-mandates (137), illustrating the real life impact of disinformation campaigns on outbreak response. When inaccurate narratives persist concerning highly technical and possibly incomplete information, there is significant opportunity for mis- and disinformation. Using partial truths, such as drawing on legitimate questions concerning laboratory safety and oversight, but connecting them to unsupported causal claims can generate very engaging conspir-

acy theories that persist even as more information accumulates to refute the claims. This can distort risk assessments by elevating implausible but dramatic events over more common and less nefarious threats, which could redirect resources to the wrong vulnerabilities. It also creates openings for malicious manipulation of the situation by state or non-state actors who may want to exploit a politically sensitive topic to sow discord or obstruct investigations.

2.7 Advances in Technology

Technology relevant for biological attribution has been advancing rapidly, especially related to sequencing capabilities (138). High-throughput sequencing methods, often called Next-Generation Sequencing (NGS), have made it feasible and increasingly affordable to sequence thousands of samples simultaneously **and allowed for complete genome coverage**. Improved methods for sample preparation, long-read sequencing, and bioinformatics allows scientists to rapidly analyze not only the genomes of pathogens but also the host and environmental DNA present in complex samples. Beyond traditional DNA sequencing, a wide array of -omics approaches are expanding ability to characterize pathogens (139). Transcriptomics, proteomics, and metabolomics are enabling a more holistic view of biological samples, identifying functional genes, protein expression profiles, and metabolic pathway changes that could suggest adaptation, engineering, or selective pressures (140).

Older methods, such as DNA profiling, relied heavily on short tandem repeat (STR) analysis and were highly dependent on intact, uncontaminated samples. These approaches offered limited sensitivity and struggled with degraded or mixed material, which is often what is found in the environment or if there is a delay in sampling (141; 142). High-throughput sequencing has overcome these barriers by directly reading the entire genome, tolerating partial degradation, and generating sufficient data to separate mixed genetic material. For attribution, this means that even challenging samples can yield meaningful data. Bioinformatics pipelines can then process massive sequence datasets to disentangle signals from multiple organisms, detect rare variants, and reconstruct evolutionary relationships.

Another significant scientific advancement lies in metagenomic sequencing, which does not rely on prior knowledge of the organism being sought. Unlike targeted PCR assays that test only for known pathogens, metagenomic approaches survey all genetic material in a sample. This enables unbiased detection of novel

or engineered agents. For attribution, this allows investigators to identify not only the pathogen responsible for an outbreak but also signatures of laboratory manipulation, such as unusual codon usage, synthetic vector sequences, or traces of cloning backbones (143). Emerging computational tools, such as machine learning algorithms, further enhance this capability by detecting genomic patterns or anomalies that may distinguish natural evolution from deliberate engineering.

The ability to analyze mixed and environmental samples has also transformed attribution. Environmental DNA (eDNA) and RNA sampling from soil, wastewater, air filters, and surfaces provides contextual information about where and how a pathogen circulates (144; 145). For example, the presence of viral sequences in market surfaces, wastewater treatment plants, or specific geographic reservoirs can indicate whether a pathogen was present in an environment prior to human cases. Attribution investigations can therefore leverage environmental sampling to determine whether a pathogen was circulating undetected before detection, or appeared abruptly, possibly suggesting a laboratory or accidental source. Key questions include whether the pathogen was previously detected in the sampled environment, whether its genetic characteristics differ from earlier or geographically distant isolates, and how its prevalence compares across locations.

Beyond sequencing, other -omics technologies play a growing role. Proteomic analysis can reveal the presence of unique protein modifications or expression signatures associated with genetic engineering (146). Metabolomics can identify unusual metabolic profiles in hosts or pathogens, pointing to selective pressures or adaptations not expected under natural conditions (147). Structural biology tools, such as cryoelectron microscopy, can be used to characterize engineered protein domains, such as those found in viral spike proteins. These methods provide a multidimensional fingerprint of a biological agent beyond simple RNA or DNA sequencing.

Equally important are advances in portable and field-deployable technologies. Hand-held sequencers, such as Oxford Nanopore's MinION, allow real-time pathogen sequencing outside traditional laboratories, supporting rapid sample processing and reducing delays in investigations. These tools are especially valuable in resource-limited or high-risk environments, where transporting samples to secure labs may be impractical. Rapid diagnostic platforms, such as CRISPR-based biosensors, **could** also contribute by enabling quick on-site screening for genetic markers of engineered pathogens **if more widely used** (123; 148).

2.8 Evidence for Attribution

Evidenced from a biological weapons investigation spans multiple domains, including genetics, epidemiology, intelligence, and forensics (traditional and microbial forensics). There should be genetic analysis of sample from victims and the environment. Not only should the pathogen be analyzed, but host and microbial communities should be assessed. Epidemiological patterns would be important to assess what could be naturally occurring compared to accidental or deliberate. Any delivery mechanism should be analyzed. Surveillance videos and social media should be included in an investigation. Intelligence should be gathered. These are some of the types of evidence expected for biological attribution. They are broad and require expertise from many different domains. Microbial forensics is concerned with collecting evidence from pathogens for investigatory purposes.

Microbial surveillance involves identifying, and sometimes characterizing, bacteria, viruses, fungi, archaea, and/or protozoans collected in a sample. Microbial surveillance programs have different goals and protocols depending on the setting and motivation. For example, during the COVID-19 pandemic, microbial surveillance is being used to monitor sewage to monitor incidence of COVID-19 in a community (149). This is an example of a public health motivated program. There are more examples of biodefense motivated surveillance programs, in part because microbial surveillance relies upon technologies that are prohibitively expensive for public health, which is chronically underfunded, unlike national security (150–152). Following the 2001 anthrax attacks in the United States (US), the US government implemented the BioWatch program to monitor the air in 30 metropolitan areas in the country for microbial agents that could pose a threat to human health (153). Intelligence Advanced Research Projects Activity (IARPA) is investing in research to detect if microbial organisms have been genetically engineered, regardless of where they are found, in their Finding Engineering-Linked Indicators (FELIX) program (154). Similarly, the Functional Genomic and Computational Assessment of Threats (Fun GCAT) program at IARPA is researching methods to advance detection, identification, and characterization of pathogens that could threaten national security (155–157). Microbial surveillance relates to microbial forensics and attribution because results from surveillance can indicate an event has occurred or become a key element in analysis.

Most microbial surveillance programs identify pathogens based on their genetic information, either by sequencing the pathogen's genome or using polymerase chain reaction (PCR) to check for the presence of

a gene an organism of interest is known to have. Programs meant to monitor for bioterrorism threats or naturally occurring threats will not be effective if limited to only one pathogen, or even several pathogens on a list (158). For bioterrorism threat reduction, there needs to be broad surveillance for anything in order to find something, because we don't know what agents a bioterrorism plot might use or the identity of the next newly emerged disease.

Notably, the evidence for biological attribution will include scientifically complex samples and analysis. Traditional forensic training does not cover all the methods that would be required to get a wholistic picture of microbial samples. For example, a key piece of evidence will likely be analyzing the genome of the microbe in question. To sufficiently do that analysis, there must be an understanding of microbial ecology and evolution. It could be easy to see a mutation in a genome and make incorrect assumptions concerning the nature of the mutation. Was it deliberately engineered into the genome or did it occur naturally? Studies might need to be conducted just to understand if the mutation has an impact on the phenotype of the microbe. These efforts could fall well outside the capabilities of a normal forensics lab, even those that regularly assesses other biological samples, such as human tissue samples.

In addition to technical concerns over conducting the investigation, there will be challenges with presenting the evidence. Most attorneys and judges are not trained to assess these types of evidence. While there are evidentiary standards in jurisdictions, such as materiality, relevance, probative force, and admissibility, it may not be clear how scientific standards compete or address those standards.

Additionally, there could be challenges with how evidence is weighed in the mind of a decision-maker. This could be particularly worrisome when people without scientific training are being asked to judge scientific evidence that appears to contradict. They could ignore technical data in favor of evidence more familiar to them, or they may rely too heavily on someone claiming to be an expert without being able to judge if that individual is, in fact, an expert.

2.9 Gaps in Biological Attribution

There have been several improvements to technology that are making biological attribution more feasible, like improved genetic sequencing and -omics technologies. However, there are several gaps in technology and implementation before the full potential of biological attribution can be met. These gaps broadly fall

into four categories: jurisdictional, political, technical, and ethical.

Attribution efforts often involve multiple agencies, including public health, intelligence, law enforcement, and defense. Without clear frameworks defining roles and responsibilities, turf wars and confusion over authority can impede a timely investigation. Before an event, different agencies need to coordinate to make these frameworks and practice implementing the frameworks in exercises. Critically, each stakeholder must understand the capabilities and primary objectives of each of the other groups. While law enforcement's primary goal will be to identify and apprehend a perpetrator, public health's primary goal will be to inhibit transmission. There could be situations where the actions of one group inhibit the work of another, if there isn't adequate planning or communication.

International attribution is hampered by limited trust between nations and geopolitical sensitivities. Countries may resist sharing data for fear of reputational harm or sanctions. The history of trying and failing to get a verification mechanism under the BWC could make it harder for countries to work together in good faith for an investigation. Realistically, contributing to an attribution investigation will be largely voluntary for many states.

While sequencing and -omics tools are powerful, challenges remain in distinguishing natural from engineered mutations, ensuring chain of custody, and integrating diverse data sources. Spoofing, or deliberate attempts to mislead attribution by disguising and agent's origin, remains a vulnerability. Resource asymmetries between national security and public health agencies also hinder the availability of cutting-edge tools to build up national attribution capacity. Technological advances must be paired with robust data management and evidentiary frameworks. Attribution requires maintaining a defensible chain of custody, reproducibility of results, and transparent bioinformatics pipelines. Standards for metadata, such as sample provenance, sequencing parameters, and computational workflows, are increasingly important to ensure that evidence can withstand scrutiny. Integration of diverse data streams requires interoperable platforms and standardized reporting to enable convergence of evidence.

Attribution raises concerns about stigmatization and mis-attribution. Prematurely labeling a country, laboratory, or community as the source of an outbreak can have serious social and diplomatic consequences. Patient zero investigations may violate privacy and increase stigma. Ethical frameworks are needed to balance transparency, accountability, and the protection of vulnerable populations.

Addressing these gaps requires integrated governance structures, sustained investment in public health and microbial forensic capacity, international agreements on evidence standards, and stronger safeguards for ethics and human rights in the attribution process.

Chapter 3

Attitudes and Expectations of Biological Investigations and Attribution

This chapter is adapted from: Warmbrod, K. L., and Gronvall, G. K. (2023). A playbook on investigations into biological events of ambiguous origin. The Nonproliferation Review, 30(1–3), 61–81.

https://doi.org/10.1080/10736700.2024.2383815

3.1 Introduction

This study sought to understand the various expectations and attitudes of relevant stakeholders toward investigations and evidence for attribution of biological events. The debate over COVID-19 origins seemed to divide people by expertise and often, the evidence that was considered when people formed their opinion was different depending on the side with which they agreed. These differences in what evidence people would consider and engage with, how they weighed the evidence, and their assessment of the evidence based on areas of expertise was the inspiration for this project. Stakeholders from a wide range of disciplines were interviewed about their thoughts on evidence for attribution. This aim presents responses from participants on both the evidence itself and the processes surrounding its collection, analysis, and interpretation.

3.2 Methods

Between February and May 2022, 41 semi-structured interviews were conducted virtually with key informants representing a wide variety of expertise relevant to biological attribution. Areas of expertise represented include microbial forensics, molecular biology, microbiology, bioinformatics, genetic engineering, virology, bacteriology, biochemistry, public health, epidemiology, international security, non-proliferation, international relations, intelligence, law, biosafety, and biosecurity. Perspective interviewees were identified based on a literature review and snowball sampling. The interviews explored participants' opinions concerning the types and characteristics of evidence potentially collected for attribution, processes possibly utilized in an investigation, feasibility of attribution, and perceived barriers and enablers to attribution. All interviews were conducted on a not-for-attribution basis due to the potentially sensitive nature of the topic and audio recorded following participant consent.

Recordings of the interview were transcribed and analyzed using a mixed-methods approach. Transcriptions of the interviews were qualitatively coded using NVivo qualitative coding software (Release 1.6.2) by an individual coder. The initial thematic coding framework was developed after the tenth interview was completed and added to as new themes emerged in subsequent interviews. The final thematic coding framework was used to code all interviews following the last interview and included categories such as investigatory steps, evidence, barriers and facilitators, and roles.

Following coding, a quantitative analysis was conducted to identify priority topics. NVivo and Microsoft Excel were used to generate quantitative metrics for all the codes in the framework to assess how frequently the topic was mentioned. Codes or co-coded pairs present in at least 10 interviews were considered priority themes for the final targeted thematic analysis. The targeted thematic analysis was conducted on the coded text corresponding to the priority themes.

This work was conducted under the purview of the Johns Hopkins Institutional Review Board as human subjects research (IRB number IRB00018728).

3.3 Results

3.3.1 Defining attribution

Participants were asked to describe what "attribution" for a biological event entailed. There was a broad range of answers; some participants considered attribution to include both determining the type of event (naturally occurring, accidental, or deliberate) and who, if anyone, was responsible for the event, while others considered only the latter question to be attribution.

Twenty-two participants noted that determining the responsible party is highly complicated. There are multiple possibilities for who could be considered responsible in different context; responsibility may lie with the person who physically released an agent, the people who created or manipulated the agent that was released, or the people who funded the work. Furthermore, other actors may have indirectly enabled a biological event, which may be enough for some observers to find them responsible. The process for determining responsibility across these different roles will vary depending on the context of the event; a funder might have unknowingly funded the development or creation of a weapon, or they may have directed money specifically for this purpose. Investigators may or may not be responsible for assigning fault following fact gathering.

3.3.2 Expectations during different stages of investigations

The methodologies and protocols used in an investigation and evidence collection were discussed with participants. While some aspects of an investigation were nearly universally mentioned by participants as critical, other areas gave rise to more divergent opinions. Considerations for attribution shared by participants were organized into groups by stage of the investigation and theme. Table 3.1 includes some of these considerations organized by stage and factor.

Mechanisms for initiating investigations

Most of the policy-oriented participants mentioned the challenges of initiating an investigation. In general, participants felt an investigation would be more plausible at the national level than the international level due to geopolitics and existing investigatory and legal structures. In the context of biological attribution,

	Stage of Investigation			
Factor	Initation	Gathering Evidence	Analyzing Evidence	Reporting
Accessibility	Host country not blocking entry War zones	Access to evidence	Techologies for analysis	All countries have access to information
Timeliness	Political willpower Public interest	Access to site(s)	Sample viability	Publicly available information
Trust	Scientist, not politician, leading investigation Potentially, no people with P5 nationality	Safety of witnesses and investigators	Validated methodologies	Accountability of investigatory team, laboratories, and involved organizations
Coordination	Creating investigatory team Travel to site	Viablility of potential samples	Triplicate analysis	Interventions
Cooperation	Between nation and investigators	Standardized methods for collection	Ownership of data	Include controls and validations
Neutrality	Geographically diverse Appropriate expertise on investigatory team	Onsite vs remote collection	Including uncertainty measurments	Just the facts, no opinons
Perception	Diversity of investigatory team Appropriate justification	Chain of custody	Between investigators and laboratories	Accuratly reflect disagreements
Flexibility	Negotiationing investigatory team membership	Logistics of travel	Between agencies involved	Privacy vs transparency
Procedures	Appropriate authority initiates Allowable scope	Ownership of samples	Between members of investigatory team	Trained spokesperson appointed and all communications filtered through them
Politics	Selection of mechanism to initiate investigation	Access to all necessary sites	Between investigators and laboratories	One designated spokesperson

Table 3.1: Matrix of factors impacting attribution and investigations of biological events identified by participants.

at the national level, law enforcement would lead the investigation in many countries. Some nations have members of law enforcement with specialized knowledge in biological events, as well as mechanisms to bring in government experts who may have additional specialized knowledge. However, many countries have never had to investigate biological events, a fact that may pose a challenge to investigating in an emergency.

At the international level, there are limited options for initiating an investigation. Two participants stated that an initiation of action under Article VI of the BWC would be equivalent to the UN secretary-general activating the UNSGM. Others stated that an Article VI of the BWC investigation could coincide with but remain separate from an UNSGM investigation. The United States, the United Kingdom, and Canada previously presented this latter view in a working paper at BWC meetings.(159) Under the UNSGM mechanism, a country may request that the UN secretary- general launch an investigation into a potential chemical- or biological-weapons event. To date, this mechanism has been used for chemical weapons investigations but not a biological weapons investigation. Two benefits of the UNSGM pathway are that the secretary-general has sole discretion to decide if an investigation will be initiated, rather than needing consensus from the UN Security Council, and there are established guidelines for how to conduct UNSGM investigations.(160; 161) However, seven participants expressed doubt about how prepared the individuals on the UNSGM roster would be if called upon to join an investigation. The list is maintained by the UN Office of Disarmament Affairs (UNODA) and currently has 530 qualified experts but is not publicly available nor is it clear how often the list is reviewed for adding new experts or removing retired or deceased experts. (162)Some participants were concerned that those on the roster may be too senior and may not work in the laboratory or field on a day-to-day basis, making them less effective or prepared to be investigators.

Another option for investigation of biological events at the international level is Article VI of the BWC, which allows a state to request that the UN Security Council investigate alleged breaches of the convention.(59) Participants were wary of this option for two reasons. This pathway relies on the Security Council, which means an investigation could be blocked by one of the permanent members. Additionally, this mechanism has been used once, in a situation that illustrated how easily the mechanism could be misused for political gain by a nation. The only use of Article VI to date was by Russia in 2022 when they accused Ukraine of creating biological weapons with the US in US funded laboratories in Ukraine. Special

sessions of the BWC were held to allow Russia to make their accusations and the US and Ukraine to respond. However, there was no formal investigation by the UN into the allegations. The Ukrainian and US governments denied the allegations and the WHO and UN said they had no evidence of biological weapons in Ukraine. Independent journalists also found no evidence. It was surmised that Russia was attempting to justify their invasion of Ukraine or distract from their own activities. This situation highlights the weakness and vulnerability of Article IV.

Nineteen participants, all of whom were scientists by training, suggested the WHO take on the role of investigating any biological events, but 20 other participants (primarily those with expertise in public health or international policy) strongly felt WHO should not be tasked with such a responsibility. Only two participants suggested the WHO could investigate a deliberate event. Most participants stated the WHO should not or could not investigate such an event, though many stated that the WHO would still be responsible for coordinating a public health response to a deliberate biological event. The WHO currently has the mandate to respond to public health events and has in the past investigated origins of outbreaks to implement public health measures to prevent further events. However, it is not within the WHO mandate to investigate known deliberate events, and some participants worried that if the WHO were to begin investigating deliberate events, it could undermine the credibility and neutrality the body needs to achieve its public health mission. If a public health event (such an outbreak) is ambiguous in source at the beginning, meaning there is no clearly identified source of the agent being transmitted to humans, the WHO likely will be the entity coordinating the response. As soon as evidence arises suggesting the event is not natural, however, the WHO no longer has jurisdiction to investigate the event's origin. Even if there is no evidence of an accidental or deliberate event, WHO must still ask permission of a member state before sending a team to investigate.

Further clarity is needed regarding who determines that the evidence is sufficient for the investigation to fall outside the WHO's mandate, as well as who would then have responsibility for pursuing the investigation. Current evidence overwhelmingly suggests the COVID-19 pandemic is a naturally occurring event. (163). Despite this, challenges arose with initiating an investigation into where SARS-CoV-2 came from, with the WHO needing to negotiate with the Chinese government to gain permission for a WHO-supported team to study the virus's origin in China; notably, the WHO was not allowed to call the study an investigation.

Forming investigatory teams and defining mission

All participants spoke about the importance of the investigatory team comprising individuals with diverse expertise and diverse geographic origins. For investigations being conducted by or on the authority of an international agency, 36 participants felt geographic diversity was crucial for the investigation team to avoid undue charges of bias. If a team were composed entirely of people from one allied group, the integrity of the team would be questioned internationally. Five participants felt that citizens of countries that are permanent members of the UN Security Council should be entirely excluded from the investigatory team to decrease perceived conflicts of interest due to national politics or avoid rejection of the investigation results based on the composition of the investigatory team.

In addition to geographic diversity, 31 participants encouraged the inclusion of diverse types of expertise on any investigatory team to provide further credibility. Commonly mentioned areas of expertise or experience were microbiology sub-specialties (the particular sub-specialty needed would vary according to the nature of the event), epidemiology, pathology, medicine, bioinformatics, biostatics, international law, biosafety, biosecurity, environmental health, social science, law enforcement, and public relations/communications. Participants also listed other areas of expertise that may be appropriate to include, depending on the situation such as expertise in a particular geopolitical region. Participants emphasized that the investigatory team should be composed of people with subject-matter expertise rather than politicians or exclusively composed of law enforcement. Three participants noted that it may be beneficial to ensure at least some members of the investigatory team actively work in the field doing sample collection, even if they are not considered to be senior experts in their discipline, due to the importance of sample collection during investigations.

Participants were asked to describe who they thought should be the head of an investigatory team. Respondents were split on whether they thought the head of the investigation should be a well-respected scientist with technical knowledge or someone with a career in international relations who is well seasoned in dealing with international politics. Notably, there was consensus that the head of investigation must be considered a neutral party and widely respected in the international community. Thirty-three participants also stated that the head of investigation should be well equipped to effectively communicate with the public, politicians and policymakers, and scientists.

Laboratories involved in investigation. Because there is no single international agency responsible for investigating biological events, especially those that may not be naturally occurring (that is, those that result from accidents or deliberate acts), there is not one laboratory that is predesignated to handle analysis of samples collected in such an investigation. Instead, laboratories whose primary purpose is something other than analyzing samples for an investigation will need to be used to assess samples the investigatory team collects. Participants were asked to describe how they thought relationships between an investigatory team and laboratories should be handled. Twenty-eight participants said the investigatory team should not be expected to conduct the analysis themselves in a new laboratory established for that sole purpose, nor should they be analyzing collected samples in their own home laboratories. Instead, laboratories that have been pre-screened, such as those on the UNSGM roster, should be used to analyze collected samples. Participants diverged, however, as to the level of autonomy the laboratories should have in conducting their analysis. Seven participants, many of whom were practicing scientists, felt laboratories should have leeway to conduct the tests they feel are most appropriate using methods they feel are most appropriate. The logic behind this is that the laboratories would be well-respected institutions and recognized leaders in their fields and, as such, would know best what should be done with samples. Conversely, 26 participants felt strongly that laboratories should conduct tests using only the methodologies ordered by the investigatory team to maintain order and ensure samples are not wasted. Three of the participants who stated the laboratory should only do ordered tests were scientists who said that by only doing analyses directed by the investigatory team, the laboratory would be less likely to be harassed or targeted for their involvement in the investigation. Fifteen participants fell somewhere in between these extremes and thought balance could be reached between ensuring the conduct of necessary tests while also allowing labs discretion to conduct additional testing.

According to the UNSGM protocols for investigations, at least three laboratories should be chosen to assess samples for any investigation. **These laboratories do not have to be accredited, just vetted by UNSGM.** Two labs would run tests as directed by the investigatory team and a third lab would receive but not test the samples unless directed by the investigatory team if the first two labs found disagreement in their results.(161) When asked if they thought the selected labs should be allowed to communicate with one another, participants again were split—although more evenly than when asked about autonomy of the labs.

About half of participants felt the labs should be allowed to communicate to optimize protocols and share information, while others felt there should be total separation of the laboratories to maintain independence in results.

Investigation Scope. Many participants, especially those with experience working in international agencies, said determining an investigation's scope and ensuring it falls within the authority of the ordering entity were vital factors. Twenty-nine participants felt that any investigation would be subject to politically motivated criticisms, as in the cases of the 2013 UNSGM investigation into allegations of chemical-weapons use in Syria and the investigation into the 2018 Salisbury Novichok poisonings.(164; 165) Most participants said that clearly defining the scope of an investigation and ensuring that the investigative body has the authority to investigate at the level of the stated scope would reduce the chances that a country could ignore or object to the investigation based on procedural grounds. One of the most frequently identified barriers to investigations and attribution was politically motivated actions to undermine, discredit, or sabotage the activity. Participants frequently emphasized the need to minimize opportunities for politically motivated attacks, especially related to the perceived integrity of the investigation or its results.

Participants often pointed to the 2013 Syria UNSGM investigation as a good example of the importance of defining scope and ensuring investigations stay within those bounds. In the 2013 Syria investigation, the mission team was instructed to investigate potential chemical weapons use but not to assess attribution due to political sensitivities. Participants familiar with the specifics of the Syria investigation described how difficult it was to stay within the prescribed scope, given how information about use of a chemical weapon could simultaneously point to who used the weapon. As an example, participants cited the difficulty in separating these two topics when examining a chemical weapon's delivery method. The Syria investigation team determined that rockets were used to deliver sarin, which is important for illustrating a chemical weapon was intentionally used against a specific population rather than being accidentally or randomly released. However, in doing a thorough investigation of the rockets, the team also conducted a trajectory analysis, which includes information that could be used for attribution of who shot the rockets and therefore used chemical weapons. Some trajectory information was included in the final report of the mission team, though participants familiar with the matter stated there was much deliberation and consternation among OPCW leadership and international diplomats concerning the topic.

Evidence gathering and handling

Participants reported evidence gathering to be a critical point in investigations and one that is particularly complex for biological events compared to others. For example, taking samples in triplicate is the standard in the UNSGM protocol, but the definition of a replicate is not as straightforward for investigating a biological event. A replicate could mean using two different swabs next to each other to collect a sample or using one swab cut into multiple pieces. Either option has weaknesses. Thirty-two participants noted there needs to be more guidance and discussion on best practices for evidence gathering, especially related to environmental sampling.

A near-universal comment from participants was the importance of maintaining evidence integrity, particularly through a strong chain of custody from collection to destruction or storage. Many participants suggested that all steps that could contribute to maintaining evidence integrity, such as filming the life cycle of the evidence or using barcodes to electronically geolocate and track samples, should be pursued. Participants often said that despite the relative lack of guidelines for biological-attribution investigations, especially considering that the technology used for such an investigation is continuously advancing, using established standards for how to handle evidence and investigations is important. Protocols used by the International Criminal Police Organization (Interpol), the US Federal Bureau of Investigation (FBI), and the UNSGM were often cited as standards that should be used for investigations even if conducted outside the respective organizations' jurisdiction.

Who collected the evidence was also widely cited as an important factor contributing to evidence quality and integrity. Samples collected by the investigatory team were commonly said to be the most trustworthy. Thirty-one participants considered samples collected by a nongovernmental organization less trustworthy than those collected by national authorities.

Types of Evidence. Participants were asked to describe the types of evidence they expected to see during an investigation. For many participants, there was a list of the types of evidence they wanted to see for any type of event, regardless of the type. Commonly listed evidence broadly falls into three categories: biological samples, witness accounts, and documentary evidence. Evidence types mentioned by participants can be found in Table 3.2.

Event-specific considerations. When asked about how the type of event could impact the evidence

Types of Evidence Mentioned
Epidemiological Data
Medical Records
Eyewitness Statements
Victim Statements
Whistleblower
Transcriptomic Analysis
Genetic Analysis of Victim
Fingerprints
Surveillance Video
Isotopic Analysis
Histopathology
Prox Card Access
Cell Phone/Wearable Tech Tracking
Cell Phone History
Search History
Financial Data
Intelligence Signals
Laboratory Order History
Environmental Sample Analysis
Genetic Analysis of Agent
Weather Records

Table 3.2: Types of evidence named by participants as important for attribution listed in order of frequency.

of the investigation, participants offered a wide range of responses. Within the deliberate-event category, participants noted there are three main types of events—assassination, war, and terrorism—with each of these event types having different challenges associated with it. For example, gaining access to the site could be exceptionally challenging for investigators in a war or conflict zone. In the case of terrorism or warfare, an actor claiming responsibility for an event could complicate an investigation. If responsibility is unclaimed, then officials may be unaware something has happened until healthcare or veterinary professionals are called upon to provide medical care to people or animals, by which point vital evidence may have been lost. In an assassination, local law enforcement officials will likely be the first people to arrive at the scene and to collect evidence, which may be a point of scrutiny (for example, suggestions of cover-ups) if an international investigation is initiated later.

Within the accidental-event category, participants mentioned two main types of events—agents accidentally being allowed to escape from a laboratory (via unintentional infection of a laboratory worker or unknown escape of an infected laboratory animal) or a stockpiling or production accident (where the accident is identified immediately but not able to be contained before affecting the wider community). Sixteen participants felt that determining if an event was accidental would be easier than determining if the event was deliberate because an accidental event should have facilities with more evidence for a positive comparison showing there was justified work with the agent in the facility, rather than a deliberate event where there may not be any samples in a facility to use for comparison. One participant felt the opposite would be true; with an accidental event, there may be significant interest in trying to cover up the event, which could drive people or institutions to destroy evidence before investigators can collect it.

Some participants suggested that all types of biological events should be considered and investigated, regardless of the initial evidence. This way, potentially relevant evidence for any scenario would be collected through the rigorous methods used by the initial investigatory team, reducing the chance of only collecting samples that could point toward one scenario over another and avoiding the introduction of bias in analyses. Additionally, 18 participants thought by starting any investigation with the understanding that the event could be natural, accidental, or deliberate, investigators would be more open-minded and more likely to find "the truth." Eight participants stated that regardless of the suspected type of biological event, investigators must assess the breadth of ramifications of the event because its implications could point to potential motives

if it were a deliberate event. Participants were very clear, however, that assessing an event's impact and potential motivations should be used to develop hypotheses that direct further investigation but that such hypotheses should never outweigh physical evidence pointing toward a different conclusion.

Access to Evidence. Thirty-three participants stated that access to information is one of the key barriers to a successful investigation. Again, using the 2013 Syria example, several participants pointed to the challenges of collecting information and samples when there is limited access to a site. In the Syrian case, the rebel groups controlling the area in which the site allocated allowed the investigatory team only a few hours to access the site safely and collect as many samples as it could. Sixteen participants discussed potential evidence that can be collected remotely, but fourteen of those participants felt that onsite evidence was stronger than remote evidence. Another consideration mentioned for investigations in conflict zones was the ability to safely interview witnesses when the group controlling the area is closely monitoring the investigators and people with whom they speak, which could sway witness responses.

Participants said timing is another critical element impacting access to evidence. Thirty-four noted how quickly evidence could be degraded or otherwise lost, especially the initial biological samples that many felt are vital for determining origin. Twenty-eight participants cited moving an investigatory team onsite as quickly as possible as a priority to increase the chance of being able to collect the most useful samples.

Evidence assessment and interpretation

Once a sample is collected, the investigatory team needs to conduct analysis and interpret results. This investigatory stage—known as evidence assessment and interpretation—elicited the most diversity in opinion from participants. Nearly all reported this stage to have a variety of significant potential pitfalls, but there were significant differences in opinion about the details of how evidence assessment and interpretation should be conducted.

In the life sciences, even slight variations in protocols can lead to drastically different results for some methodologies.(166; 167) There are also often several slight modifications that could be made to a protocol, any of which can produce a slightly different result. For example, there are **dozens** of protocols to sequence one pathogen. While this is helpful for generating data for scientific exploration, it is a challenge for an investigation. Just as participants cited chain of custody as critical for evidence integrity, many said having

validated and trusted methodologies is similarly vital for integrity. Twenty-seven participants felt that an investigation should stick to methods that are well known and widely used, as these are methods that are more easily assessed by and familiar to the broader scientific community and easier to validate. Based on these results, using tools and methods that are more understandable and accessible may increase the willingness of policy makers to accept results and decrease opportunities for critics to attack an investigation. Alternatively, 12 participants said using cutting-edge methods may improve an investigation because they could provide highly relevant information that more traditional methods could not.

To increase trust and validity, twenty-nine participants remarked that having guidelines on the conduct of evidence assessment and interpretation would be helpful. However, opinions diverged over the content of such guidelines or standards. One participant stated:

"We have to be careful when talking about guidelines. We need to be specific about the type of testing that is going to be appropriate under different circumstances, without being so specific to say what you must do. Because number 1, that will timestamp our recommendations and not allow innovation to happen if something more effective comes along. Number 2, it would undermine those who are leading the investigation, and who will be working in the labs to use their best judgment of what is going to be the most appropriate under the circumstances."

Thirteen participants stated that gold-standard methodologies should be identified and any investigation should prioritize these methodologies over others, with some also suggesting further methodologies (such as those on the cutting edge) could be used alongside the gold-standard methods. Six participants felt no cutting-edge methodologies should be used because they might risk producing results that are not well understood, creating room for questions and doubts. Eight participants suggested that instead of listing methods that should or should not be used in an investigation, having the international community agree upon and create parameters for selecting or modifying protocols for evidence analysis prior to an investigation would be useful. Additionally, 33 participants said the development of guidelines on how to demonstrate and report validation of methodologies, including appropriate controls, would be helpful.

In addition to the breadth of protocols for sequencing and other evidence-analysis steps, there exist many options for subsequent bioinformatic analysis. Typically, scientists conducting analyses must decide on the parameters to include in a model, values for those parameters, methods of cleaning data, and many

other issues. Changes to any of these aspects can produce different results. Some participants highlighted the diversity of choices available for analyzing data as a key area needing assessment and said decisions on these methods prior to the start of an investigation could alleviate wide disagreement concerning selection of methods, even among the scientific community. Nine participants expressed a desire reach more clarity and agreement on which parameters, models, and databases could be used in an investigation prior to initiating an investigation. There were concerns about data security if an investigation used previously published information, such as using common genome repositories for assessing possible natural origins of an agent. Such previously published data could have been intentionally planted or falsified. Thirty participants noted that the sensitive nature of an attribution investigation may be leveraged by some people looking to discredit the process; therefore, it is of utmost importance that those wishing to discredit the investigation are unable to point to the investigation's science to achieve their goals. Highlighting the importance of appropriate analysis and reporting of results, one participant stated, "If the science is wrong, if the science is reported misleadingly, if the report overstates the science, the investigation is dead."

Weighing evidence. Understanding how people weigh different types of evidence and how that weighting affects interpretation and evidentiary conclusions is of great importance for informing future investigations of biological events. Not all evidence is equally informative; direct evidence is stronger than circumstantial evidence. Evidence that commonly is part of criminal investigations and trials, such as fingerprints or surveillance footage, may be more readily accepted and understood by the public than evidence collected using more unfamiliar and complex methods. This could lead to more familiar types of evidence being more heavily weighted, even if they constitute less direct evidence.

Participants were asked what factors they considered when comparing different types of evidence. Thirty-five participants could not cite inherent qualities that made them value one type of evidence, such as biological samples, over another, such as surveillance footage. Rather, participants named the source of the information or evidence as the top characteristic, followed by validation (controls used in tests, proof the methods were validated against other methods, etc.). The source of the evidence often was described as the most important attribute because "if the source isn't broadly trustworthy, it's junk." Nearly all participants who specifically mentioned trustworthiness of sources/collectors of evidence during their interview (a total of 17) ranked them from high to low trustworthiness as follows: investigatory team, United Nations agency,

well-respected international nongovernmental organization, national health or public health agency, local health-care provider or clinic, military, national law enforcement, and local police. The first four sources, in the order listed here, were the same for all 17 participants who offered a ranking of sources. For 14 of the 17 participants, national health or public health agency ranked above military, national law enforcement, and local police. For the remaining three participants, national-level authorities or organizations ranked above local ones.

Fifteen participants discussed difficulties with trusting evidence collected from law enforcement compared to public health or medical providers. While law enforcement was widely considered to have the experience and expertise to collect evidence, some participants stated they were less likely to trust evidence from law enforcement or the military because such organizations were perceived to be more likely to be biased than public health or medical practitioners. Additionally, nine participants said their ranking of sources' trustworthiness might change depending on the type of evidence or country in question. One participant said ranking sources for trustworthiness was impossible because different sources potentially could offer different types of evidence, and the trustworthiness of those evidence types could not be meaningfully separated from the trustworthiness of the source.

Thirty-five participants said they would be most convinced by a convergence of several different evidence types. For example, participants found that the strongest case for evidence results when multiple -omics methods, video surveillance, and intelligence signals all support the same conclusion. Participants were split on whether they would trust a conclusion drawn entirely from circumstantial evidence; 22 felt there must be biological samples to support a conclusion while others said there may be situations in which they would have to trust results based entirely on circumstantial evidence because that may be all that is available. However, for those in the latter group, many stated that epidemiological and medical data, at a minimum, would be necessary if no biological samples were available.

When asked if there was a single piece of evidence that would be weighted most heavily, eight respondents said that a definitive delivery mechanism would be the most important evidence to them, as it indicates not only the occurrence of a biological event but also the intent behind it. However, no participant felt that identification of a delivery mechanism alone could be used to conclude an event was deliberate in nature or who (if anyone) was responsible for the event; the respondents said they would still need to see several

pieces of evidence supporting that conclusion to believe that the event was deliberate.

Other than a delivery mechanism, the only other type of evidence participants named as most impactful or important was genomic evidence. Thirty-two considered genome comparisons vital for biological events. One participant stated, "The only way to definitively state beyond a reasonable doubt that something is naturally occurring would be to find the exact genetic match in an animal somewhere. Even then, I would need to see evidence for how it jumped to humans. It's not enough on its own, but it is the single most important piece if it can be done." Phylogenetic analysis for tracing transmission or inferring evolutionary history was considered by some to be less impactful than direct comparisons of genomes. One participant noted that more models and assumptions are needed for phylogenetic analysis compared to matching genomes, thus weakening the level of certainty that could be derived from a phylogenetic analysis, which is why they considered the genome matching the most impactful analysis.

Addressing uncertainty and disagreement. Unlike other WMD, biological agents can evolve and are ubiquitous in the environment with dynamic populations. Such characteristics mean that one sample taken at one time may not be identical to another sample taken in the future. Additionally, evolution is not predictable, and our current technologies are not capable of continuously monitoring or measuring microbial populations at the organism level. The rapidity of the changes, the limited resolution available to measure changes, and the inability to predict changes mean there are several ways in which uncertainty may be introduced into biological measurements.

In addition to the uncertainty inherent in biology, there is uncertainty introduced through the methodologies for sample analysis (both physical lab methods and bioinformatics methods). The -omics methods likely to be used for attribution rely on statistics and report likelihood scores or confidence intervals or both, as these methods rely on assumptions built into the statistical method. Such uncertainty has complicated the use of genetic information in courtrooms in other types of cases, particularly those relying on a single sample of DNA from a crime scene compared to a suspects DNA.(168)

Participants agreed that, where possible, an investigatory team should take steps to decrease the uncertainty introduced in the investigation and be transparent about where uncertainty does exist. However, participants diverged in how they thought this would best be accomplished. Fifteen participants, primarily those with training in a scientific discipline, said an investigatory team should quantify uncertainty within

their investigation and results by reporting a number that expresses certainty in their result. Twelve participants, primarily those whose expertise was in policy or law, said simply explaining where and why uncertainty exists would be sufficient to avoid unnecessary confusion. One participant went so far as to say that reporting a specific number for certainty would "be the downfall of the entire investigation." The scientific literacy of the audience was of concern to 33 participants, as many politicians who would be making decisions about what to do with the results of the investigation are not trained in science or statistics. They were concerned that a report too heavy with technical, scientific, or statistical detail would lead to confusion and make decision makers less likely to act on the results, even if given an executive summary written for a non-expert audience.

Reporting results

Participants frequently discussed the importance of having a strong strategy for communications with both politicians and the public. Many participants mentioned the importance of not leaking information and speaking to the media only if authorized by the head of mission, as there needs to be very clear and structured messaging surrounding the investigation. Twenty-five participants warned that an attribution investigation is likely to spark mis- and disinformation from those with political or economic motivations to undermine the results of the investigation. It was suggested that to combat the spread of false narratives, an investigatory team would need to include a highly trained spokesperson capable of clearly discussing technical details while also navigating the political environment. One participant felt this spokesperson role was the single most important position in the entire investigation.

Many participants cited the controversy over the origins of COVID-19 as an example of the enormous amount of mis- and disinformation that may circulate in response to an investigation or attribution. Twenty-seven participants felt any biological event, regardless of the type, origin of the agent, or amount of evidence, would be subject to at least some conspiracy theories, in part because of how intimately health threats affect individuals and societies. Fifteen participants who discussed mis- and disinformation often described a phenomenon in which the amount of evidence publicly available would inversely correlate to the quantity of mis- and disinformation and conspiracy theories; events with little evidence—particularly those that most experts believe to be naturally occurring—were speculated to be highly vulnerable to disinformation.

To prove something is not deliberate or accidental in origin, and thus naturally occurring, likely would require proving a negative. Alternatively, twelve other participants speculated that events that have large amounts of evidence would be the most highly susceptible to an increase in conspiracy theories and mis- and disinformation. Events with plentiful evidence were believed to be vulnerable because those so inclined could create more narratives to promote false contradictory evidence or otherwise poke holes in official stories. While no participant suggested limiting the amount of evidence made available to the public, twenty-six expressed extreme concern over the potential misuse of evidence for such purposes and stressed the importance of a communications plan with considerations for dealing with mis- and disinformation.

Fourteen participants said that policy, social-science, and communications experts on an investigatory team should be involved into writing the report. These participants noted that scientists often are trained to write for an audience of other scientists, and that is not the sole audience of an investigation report. There was a division, primarily along disciplinary lines, in the responses to the question of how best to report results. People who identified as scientists often said reporting results should include a quantification of uncertainty or confidence in results, such as stating an exact percentage of confidence in different hypotheses covered in the report. People who self-identified as policy, legal, or international-relations experts often said that having specific confidence numbers would create serious problems for those needing to act based on the report. These experts said instead of exact numbers, the report should be clear about which results are ambiguous and should use language that is clear about what is known or not known, without overstating what the evidence and subsequent analysis concludes.

Thirty-four participants felt it would not be possible to conclusively state the exact origins of an agent causing a biological event. Rather, the goal should be a "convergence and preponderance of evidence that paints a picture" of what happened. There should not be an expectation that every piece of evidence will prove a theory or that any one piece of evidence alone will prove a theory. Instead, all the evidence should be considered together. There was wide acknowledgment among participants that interpreting evidence would be a challenging, "murky" process as many pieces of evidence, if considered individually, could be interpreted to support multiple different hypotheses.

People seeking to discredit the investigation or to support a specific agenda are likely to cherry-pick evidence. The investigatory team cannot be seen as also cherry-picking evidence. This drives the need for clear,

transparent, and fact-based reporting of the results. Some participants said there could be a temptation for an investigatory team to present only evidence that supports their proposed narrative, but these participants emphasized that evidence supporting all theories should be collected, analyzed, and reported.

3.3.3 Technologies to enable attribution

While there was no consensus regarding the likelihood of an investigation being successful in identifying a perpetrator, most participants felt the potential for success of attribution was increasing with technological advancements. In fact, many participants felt that technology was not a limiting factor for attribution; instead, legal, political, and social factors were considered greater barriers. One participant stated, "Science has advanced rapidly, and we are ready, but politics, laws, and policy hold us back." High-throughput -omics methods, such as genomics, proteomics, and transcriptomics, and computational advancements were the most mentioned technologies believed to contribute to increasing feasibility of attribution.

Genetic technologies

Genetic sequencing was widely reported as vital for attribution of microbial agents. These results are important because by using sequencing, investigators can identify the pathogen and look for signals that could point to origins of the agent. Having the genetic sequence of an agent enables investigators to make comparisons between various samples collected as part of the investigation (such as samples collected from different victims or locations) or between a sample in evidence and sequences in repositories (such as Genbank or Addgene). Such comparisons are particularly helpful for assessing whether a common agent is causing disease, comparing evidential samples to environmental samples to assess the possibility of natural occurrence, and comparing samples gathered from victims to samples collected from laboratory or other spaces being investigated as potential origins. Genetic sequences also can be analyzed to consider genetic signatures that may indicate whether the agent was genetically manipulated; help delineate between a naturally occurring, accidental, or deliberate event; or provide information that may indicate if a particular laboratory supplied or manipulated the agent. Additionally, if the biological event is causing disease in multiple individuals, information on the pathogen's genetic sequencing can be used for molecular epidemiology to better understand transmission patterns and evolutionary history, both of which could point to attribution. **Quantification of**

genetic material can also be useful for attribution by indicating if there was trace amounts or more significant concentrations of a pathogen present, which could point to deliberate activity.

Thirteen participants also discussed the limitations of using genetic sequences for attribution. Genetic sequencing is useful for identifying pathogens based on differences in DNA or RNA genomes but is not particularly useful for understanding functional differences in how those genes are expressed, which are the basis for pathogenesis and other signals that are used to monitor and categorize an event. Seven participants mentioned that genetic sequencing often is cited as the primary data needed for attribution, but in practice, sequence data alone are insufficient for attribution. While the genetic sequence can be used to infer helpful information, such data have limited ability to demonstrate whether any intent underlies an event. Even if an agent were determined to be genetically modified, additional information would be needed to understand the motivation leading to such modification.

Participants were split about whether environmental surveillance for microbial agents would produce useful information for attribution. Environmental surveillance entails the broad collection of microbial agents in a given environment and the cataloguing of the results. Seven participants felt catalogued environmental samples could be compared with samples collected from an investigation site, to possibly identify a geographic location of origin. One participant mentioned that the practice of environmental surveillance alone could be useful as it could act as a deterrent; even if the information produced is not useful in an investigation, if potential adversaries think environmental surveillance produces useful information, it could decrease their interest in pursuing a bioweapon. Ten participants felt that environmental surveillance would not be useful for attribution because investigators would have to be sampling constantly from all locations to have sufficient data to track an agent's origin. They argued that because biological agents evolve and microbial population structures can change frequently, one-time collection would have limited utility for attribution.

Other -omics-based technologies

While genetic sequencing was the single most common technology reported as important for attribution, other -omics-based approaches, particularly transcriptomics and proteomics, were also regularly mentioned. Because these methods assess molecules that are directly involved in function of the agent, they can provide

different information than genetic sequencing. Proteomics, for example, can provide signatures that indicate different pathways of pathogenesis or virulence between different strains (169). In roughly the same way that genetic sequences can be used as comparators, protein signatures may also shed light on where or how an agent was created or manipulated, such as if it was cultured in a certain animal or cell line (170). Proteomic signatures may also help point toward intent or motivation, more so than genetic signatures, as proteomic signatures refer to a function or characteristic directly. In addition to proteomic signatures to assess identity or unique functions, four participants also mentioned using mass spectrometry for isotope analysis, which could provide information about the geographic location where an agent was created.

Computational tools

Participants also frequently suggested computational tools for advancing attribution. Some participants mentioned the use of algorithms designed to identify potential source laboratories but were skeptical this approach would be feasible. Other computational tools that were suggested included programs that would comb through computer or institutional purchasing histories to look for evidence of malicious intent or abnormal behavior, programs to analyze social-media histories looking for evidence of past locations or possible motivations, or programs that combine different types of data sources to assess behavioral patterns. Like the other technologies, computational tools were considered to have limitations. The most common one mentioned was the reliance on quality data needed to develop and use such algorithms. There was doubt among five participants that the repositories used to collect sequence information would be sufficiently reliable for attribution, given the relatively small sample size of data included in the repositories and lack of security measures protecting many public repositories. For other data sources, a few participants noted concerns about the depth of surveillance feeding such algorithms and the potential for invasion of privacy.

3.4 Discussion

This research was inspired by the diversity of viewpoints surrounding the origin of COVID-19 debate; experts frequently and strongly disagreed with each other over the validity of evidence, how it was collected, and even what counted as evidence. In many cases, it appeared that experts in microbiology, especially virology, were using genetic and phylogenetic analysis of environmental samples, early patient samples, and

samples collected animals in making their determination that COVID-19 was naturally occurring. Proponents of an accidental or deliberate origin for COVID-19 were often not virologists and came from other scientific disciplines or were not scientists at all. This group frequently referenced the proximity of a laboratory doing research on coronaviruses in China to the first cases of COVID-19 and spots in the genome that they said could be evidence of genetic engineering of the virus. The divergence in conclusions regarding COVID-19 origins between specialties, especially in how these specialties were weighing different pieces of evidence, inspired the questions asked of interviewees and the breadth of experts that were interviewed. Unsurprisingly, the results from interviews conducted here in many ways mirrored the divergence across disciplines seen with the origins debate for COVID-19.

Sample collection and analysis were the two areas with the most divergence among interviewees, which often stemmed from competing demands between rigorous science and expectations of an investigation. While everyone agreed that the science used in an investigation must be sound, the balance between best scientific practices and best investigative practices varied drastically between different disciplines. Differences between interviewees often followed a pattern of scientist versus security or public policy expertise. Scientists most frequently advocated for laboratories to have leeway in choosing how they analyze a sample, using whatever tools were available to analyze samples, and including more detailed information in reports. Non-scientists consistently advocated for less autonomy of laboratories, using fewer methodologies, and less technical information included in reports. In an attribution investigation, the investigatory team will need to understand this dichotomy when designing their investigation. Messaging to the scientific community about concerns for investigation validity and messaging for non-scientists about rigors of the scientific method, validation strategies, and methodology strengths and weaknesses may help address competing concerns during an investigation and ultimately help people accept results without falling prey to mis- or dis-information.

Scientific training involves a heavy focus on examining appropriate methodologies for addressing a question. Scientists often innovate on prior methods to better study their topics or use their samples. Methodologies undergo rigorous validation through repeated testing and peer review. Especially in biology, where samples can be very different, an iterative process is often helpful to optimize methods for analysis. For those unfamiliar with what the validation process entails, it may appear that changes to protocols or new

methodologies could be short-cuts. They may prefer methods that are older that they may remember hearing about from a science class long ago or one that is more basic and easier to understand without expertise in the field. Such methods may be easier to trust for those not working in that field. In this work, participants frequently talked about trusting science and the investigation process, but much less frequently did that trust apply to the scientists as well. By focusing on the methodologies and protocols, these participants could be avoiding needing to trust the scientists. Trust in scientists and people with traditional academic training has eroded among the public since the COVID-19 pandemic. Non-scientist participants could be reflecting this societal shift away from trusting science. Alternatively, biology is a field with complex, evolving systems. There is an inherit amount of uncertainty in biology. People who are not trained to understand and accept that these are dynamic systems can be uncomfortable with not having definitive answers immediately, which could contribute to a willingness to accept something like geographic location as a strong indicator of origins over genetic analysis.

Sample collection will be a very challenging part of any investigation since everything downstream relies on good sample collection. The UNSGM guidelines don't have jurisdiction over non-UNSGM investigations. The FBI has their own protocols for how to conduct sample collection and they are not identical to the UNSGM guidelines. The UNSGM has guidelines for how to collect samples, such as taking all samples in triplicate, but even that has limitations. Given the nature of microorganisms, taking samples in triplicate isn't perfect. If trying to assess the microbial population composition on a surface, you could use three swabs side by side to get three different samples from the same general area. However, each swab will cover a slightly different area, and the microbial populations won't perfectly match. Alternatively, one swab could be used and then the sample from one swab split into three. However, there could be a low concentration of the microbe on the swab and splitting it into three could result in the microbe being below the limit of detection for a test in all three split samples. There isn't a clear answer for how to take replicates of samples that meet standards for both scientific rigor and evidentiary gathering for law enforcement. The backgrounds of those making decisions based on evidence will likely determine how willing they are to accept and heavily way nuanced data that has a confidence interval. Those looking for a black and white solution may not find it in biological data.

Once samples are collected, determining what methodology to use to test the samples will also not be

clear in all cases. While there was clearly a difference in opinion among responses in these interviews, determining which methodologies to use during an investigation will rely heavily on the context. For a high visibility, international investigation, the priority may be to convince heads of state internationally of the results. In this case, using gold standard methodologies over newer, less well known but possibly more powerful tests may be the best choice since gold standard methods may be more trusted and familiar globally. Alternatively, if the US is investigating an event and only the President and Congress need to be convinced of results of the investigation in order to act, then investigators likely have more leeway to approach researchers developing cutting edge methodologies, especially for methods developed at institutions in the US. Investigators will need to assess the context of the investigation, especially considering who the primary audience of the report will be and who will be responsible for taking action after based on the results, when determining the best course of action concerning sample processing.

Any international investigation will be a huge endeavor across many laboratories and scientists. There will be several steps taken to maintain chain of custody and evidence admissibility in court. However, these steps will take time and likely require samples to be shipped to different areas of the world. Not only will logistics of planning the movement of samples being challenging, but preserving the samples themselves in transit and within laboratories will be a critical step. Many biological samples are sensitive and samples may be altered by transportation. A sample going to Asia could be affected differently than a sample sent to Europe based on the conditions and time of travel. Additionally, samples may be processed at different speeds or under slightly different conditions once in the laboratories, which could affect results. Reagents, equipment, and slight variations in researcher's techniques could differ and further lead to differences in analysis. Validating across laboratories will be another massive task. There could be ways to minimize differences between investigators, such as requiring every laboratory to use the same reagents from the same suppliers, using robots to do sample preparation, or having all researchers come to one location but work independently, but there will still be opportunities for small differences. Part of the investigatory team's job will need to be determining how much difference of results could be expected from failures in sample preservation of validation across labs. There must be a balance between the need uniformity across labs and what is possible within budgets, logistics, and available resources.

Following COVID-19, there has been a lot of discourse around environmental sampling and cataloging

microbial populations outside of an investigation or before an investigation is needed. Some, especially those without microbiology expertise, have stated that environmental sampling poses too high of a risk to public health. They worry that by sampling more microbes, there is a higher chance of accidental infections in the laboratory or during sampling, which could then start a new pandemic. However, environmental sampling has been done for decades with low risk to researchers or the public. Additionally, the wealth of information that could be gained from analyzing such samples can help prepare for future outbreaks and enable better therapeutic development for different infections. Environmental sampling, which was not a contentious topic before COVID-19, has unfortunately been roped into the gain of function research debate, which is another area where people without expertise in virology or microbial genetics are often trying to tell people with expertise in those areas what research they can or can't conduct.

In many cases, the first sign that something is amiss is disease presentation to a health care provider. The initial samples for the event may be the most informative for determining who is responsible, but they are also the least likely to be collected by the investigatory team. Samples collected by medical practitioners or public health officials are collected using protocols established with the aims of those organizations in mind, rather than law enforcement or investigatory goals. The chain of custody likely would not be as stringent as the standard used for law enforcement. Additionally, samples taken for public health purposes may be anonymized to decrease the risk of privacy invasion, but this would impede an investigation. Any investigatory team will have to reconcile their need for certain samples with the limitations of the sample source.

There will be jurisdictional concerns in the event of an attack. A country is not going to want to give up its right to investigate an event within its own borders or to be forced to work in tandem with an international agency without control over the situation. Events that affect citizens of multiple jurisdictions could create situations where countries do not want to share information or samples with the other countries, even if that limits the investigation. There have been examples of jurisdictional issues between agencies within the same country inhibiting non-biological event related investigations, such as different parts of the intelligence community investigating al-Qaeda before September 11, 2001 and not sharing their information with one another. Biological events are more rare than other types of naturally occurring or human-influenced events and often more complex, which is likely to exacerbate complications with jurisdictional issues.

Any investigation will have to develop a report of the findings. There will be different philosophies on how to do so; create something simple that is easily understood by a wide audience but lacks nuance and detail, create something highly technical that includes all data but could be misinterpreted by non-experts, or find a middle ground between these two extremes. Which is most appropriate will change by situation. At no time should information be included that is inaccurate, but there may be times when a slightly simplified result is included. Investigators must know their audience and understand what the decision-makers expect as a deliverable. Will the decision makers be utilizing an entirely independent group experts to interpret results or will the decision makers themselves be interpreting? In an ideal world, scientists would be able to write the report with all results and detail that would normally be included in the report and those responsible for interpreting the report and making decisions based on the report would understand it in that format. In the real world, scientific literacy is low and many decision makers are willing to use their version of "science" to justify decisions that are not supported by science. An investigatory team will need to be adept and skilled in understanding the broader societal context in which they are working and find ways to report truthful results that are not easily misrepresented.

In most scientific endeavors, the scientific community champions being inclusive with data. While publishing often does not support a system of reporting non-significant or negative results, such results are important for understanding the systems being studied. There is movement toward changing the current publishing paradigm to be more supportive of such results to be more aligned with the goals of science. In this vein, including all results, even those that are not easily interpretable or fitting of a given narrative, should be included for an investigation report. Determining error rates and confidence intervals from each type of test or across laboratories or methodologies should be included in the report. Any results coming out of laboratories should include all information about methods and tools used, results from controls, raw data, and accounting for the original sample itself over its lifetime in the laboratory. The final investigation report should indicate that all details included from the laboratories were reviewed, but doesn't need to have raw data or all of details about controls and validating tools. Instead, the investigation report should focus on analyzing results from across laboratories and interpreting their meaning.

For example, if there is an investigation where three labs analyze an environmental sample believed to contain a pathogen of concern, then the investigatory team should receive results from all three laboratories.

Those results should include details about the exact protocols they used, lot numbers for reagents, exact recipes for any media, name and proof of validation of machines used (like a sequencer), and the raw data from samples and controls (such as sequence reads or pictures of growth plates). The investigatory team should review all information given, request any further information that may be missing, and then compare across laboratories. The investigatory team report should specify the type of information they received and indicate where there were differences between labs. The investigatory team report should include the results of the tests across the laboratories without relying on all raw data. A list of what was found in the sequencing across samples would be appropriate whereas including all sequences themselves would not be appropriate. It could be appropriate to make the raw data available beyond the investigatory team in some context (there could be privacy issues if human DNA is in any sample), but there will likely be too much and too technical of information to include all information in a report meant for politicians. It may also be appropriate for the investigatory report to include some interpretation or contextualization of results, such as indicating if results were abnormal for the environment from which they were sampled.

The investigatory report should include some indication of confidence in results. Presumably, the investigatory team includes experts. There must be some level of trust in their ability and knowledge to make judgment calls relating to their area of expertise and gage how much they believe in the results. The exact method of reporting confidence could change with audience. The report should clearly delineate between results with low level of confidence and results where the investigatory team is not confident in the interpretation of the result. Lack of being easily interpretable should not automatically mean that result or evidence should be ignored or weighed less that other evidence, but it should be noted where there is not a clear interpretation or disagreement among investigators on the interpretation. People making decisions will need to be able to handle uncertainty or ambiguity in results either for themselves or based on trusted counsel, which will require some science literacy.

Scientists will be needed for a valid investigation. While some scientists may be uncomfortable being in a position where they are involved in a contentious or political practice, there must be some that are willing and trained to do so. There also needs to be an effort to rebuild trust among the public and politicians in science. Without more trust, an investigation may be ignored or be limited to evidence gathered by other sources that are more trusted in the eyes of the public, which may lead to incorrect conclusions. The stakes of

an origins investigation will be high because the potential consequences for someone at fault, or perceived to be at fault, can range from prison to potentially starting a war. Russia attempted to justify invading Ukraine on allegations of biological weapon production in Ukraine. Beyond geopolitics, there will be some who seek to use origins to limit science or push for other far-reaching policies that could negatively affect societies.

Part of the importance of picking the right mission leader is picking someone who knows their audience and the science. A scientist with limited experience working with politicians or policy makers without a science background may be inclined to go with the cutting-edge methodologies if they feel this will produce better results. However, many politicians and policy makers rely on advisors who may have some understanding of the science but are not as familiar with the most up to date methods. If these advisors don't trust the methods used, that will impact the politicians. The mission/investigation leader needs to know who they are targeting with their report. Will the politicians getting this report have advisors familiar or open to new technologies or are they more traditional? In international relations and politics, networking and understanding who you are talking to is vital.

3.5 Conclusions

Results from the interviews conducted here highlight the diversity of opinions surrounding attribution and investigations of biological events. Beliefs that are strongly held and considered obvious to someone with one background sometimes were directly contradictory to beliefs held by others with different backgrounds, such as on the question of whether an investigatory team should use standard methods exclusively or also use more cutting-edge technology. Results from this research demonstrate why effort must be put into developing consensus and clarity on protocols for attribution before an event happens.

Identifying the origins of biological agents responsible for threatening human, animal, or environmental health is critical for developing and implementing effective interventions to prevent repeat events. Developing clear guidelines and tools for investigation and attribution before they are needed is vital for their successful use during an emergency. Notably, however, efforts to identify a biological agent's origin likely will always be a sensitive exercise. There will be many nuances specific to each scenario, and those responsible for responding must prepare for uncertainty and clarity, especially when discussing results with the public or non-scientists. The nature of biology means there will always be some uncertainty when analyzing

a biological event. This uncertainty can be minimized, but it will never be eliminated. Investigatory teams will have to recognize and reckon with uncertainty from many sources over the course of the investigation. If uncertainty is not adequately addressed or represented to decision makers, many of whom probably will not be scientists, could have long lasting consequences for the investigation and world politics. Stakeholder attitudes and opinions will evolve over time, whether over the course of a single investigation or between events. Flexibility, adaptation, and scientific literacy will be vital characteristics for all stakeholders, including investigators, those responsible for acting on the results of the investigation, and the public.

There must be greater clarity on what policies and pathways are available for international investigations. Clarifying what agencies have authority to investigate different types of events, especially those that have ambiguous sources, is an immediate priority. The WHO, in particular, requires clarity about the point at which it is not responsible for investigating a disease outbreak and at what point another entity, and which one, will take over an investigation. Clarifying who has the authority and responsibility to conduct investigations is not only a logistical imperative. Leading or even participating in an investigation can leave an organization vulnerable to mis- and disinformation and other attacks that could degrade public trust in the organization. For entities like the WHO, whose mission in the public health space is dependent on trust from governments and the public, participating in an investigation has potentially grave consequences that could impede its ability to fulfill its mission. Certain nongovernmental organizations that provide health care or other necessities could similarly be blocked from continuing to provide aid if they participate in an investigation or provide samples or evidence. Such groups require more clarity on what they are legally obligated to do to develop their own policies and procedures for investigations.

There also should be an international effort to gain consensus on methodologies that would be acceptable in an investigation. Even without consensus on all types of methods that could be used, identifying gold-standard methods that could be used to validate new methods would be useful. When seeking such consensus, there must be representation from multiple disciplines and nationalities for the results of such work to be seen as acceptable and fair. It is critical to ensure that any established guidelines on analysis techniques and investigatory protocols are accepted and validated across stakeholders and regions. The existence of broadly endorsed or accepted standards is likely to blunt criticism levied by a country or other entity against an investigation on the basis of methods or conduct. Additionally, standards can help an

investigatory team make decisions, rather than having to guess acceptable or best options. Additionally, guidelines about communications strategies, including how to manage mis- and disinformation, are important. As new technologies are developed and become available, stakeholders should regularly review them to determine if they warrant any alterations in protocol.

The science and technology of biological-event attribution investigations are solid, but policy, political, and implementation barriers are inhibiting this work. Understanding expectations for evidence and the investigations to collect evidence is the first step in creating a usable and viable attribution framework. Developing a playbook for attribution could help identify areas of uncertainty and necessary research, such as assessing how different types of evidence are interpreted and how the factors identified here may influence decision making and action following an investigation.

Chapter 4

Assessing Evidence from Biological Events in Fictional Scenarios

4.1 Introduction

In any investigation into a biological event, there will be many different types of evidence, some of which will probably involve complex genetic analysis. Different people will use collected evidence differently when drawing conclusions. The goal of this project was to examine how different people assess and interpret evidence for attribution. What evidence is considered more impactful for decision making compared to others? What factors could impact that trustworthiness? Results from aim 1 provided the types of evidence people would expect to see for attribution, but given incomplete and imperfect information, how is the evidence available considered?

Fictious scenarios and evidence were developed for participants to consider. Scenarios were chosen to ensure each participant had the same information in the same format and to make the exercise less theoretical. Scenarios were fictious to avoid participants having prior knowledge of the event. Two scenarios were devised for this exercise. They were designed to not have one clear answer and the evidence included in each scenario was selected to leave uncertainties. The goal was not to have one "right" answer and see who got it right, but instead to assess how people used different and possibly contradictory evidence when there wasn't a clear answer.

Scenario 1 is broken up into 2 parts: did something deliberate occur using a biological weapon and can we determine who did it. This scenario is meant to be more focused and have a specific, identifiable event in question. Follows a more well recognized investigatory pathways (murder or possible terrorism). Scenario 2 is meant to be more nebulous and represent a situation where an ongoing event with an unknown origin is being investigated, such as the beginning of a naturally occurring pandemic.

Participants were asked to complete a survey following reading each scenario. The survey asked for evidence provided to be ranked, perceived credibility of the evidence, and confidence in conclusions.

4.2 Methods

4.2.1 Scenario Development

Two scenarios were designed for this exercise, first by writing a prompt of what "really happened" and then creating evidence that may be collected for such an event. An assassination was chosen as the focus for one scenario because it was more similar to a traditional police investigation that many people would be familiar with to some extent, through personal experience or pop culture. A widespread change in health status scenario was meant to reflect a biological event of unknown origin and type, such as the early days of a pandemic from an unknown pathogen or a long-term contamination of the environment affecting health. The evidence selected for inclusion in the scenarios was chosen based on responses from experts interviewed in the work completed in Chapter 3.

The goal for writing the evidence in each scenario was that the evidence be at least partially understandable to a broad audience, but to avoid biasing participants by over-explaining what the evidence meant. To do this, we stated what the evidence was, how it was gathered, and the fact of what it showed. For example, in Scenario 1, the "Samples collected from the IOM rabies lab" evidence was reported as, "the highest nucleotide sequence similarity between any sample from the rabies lab and the Salle samples was 83.2." The percent similarity between genomes was reported but there was no context provided for how this percentage compared to other rabies genomes.

Limiting the evidence to ten briefly described pieces of evidence per section was done for feasibility of this study and to represent high level findings that may be reported to decision makers based on a more

thorough investigation report. Participants were asked to interpret results provided to them rather than to act as investigators themselves.

Scenario 1

Scenario 1 focuses on the death of a rich CEO named James Salle who is trying to build an unpopular oil pipeline through a foreign country. The entirety of Scenario 1 is included in the appendix.

This scenario was sectioned into two parts; the first asked participants to determine if Salle was deliberately killed with a bioweapon and the second asked participants to assume he was intentionally killed and determine who threw the dart that hit him. Rabies was chosen as the biological agent because it has a clear natural transmission pathway, and it has a relatively low incidence of human cases making it less probable that he picked it up outside of the event described in the scenario. Rabies also has a long enough incubation time that an infection could plausibly be missed for weeks, delaying the start of collecting biological samples. Finally, rabies is an RNA virus that can make a quasispecies, which makes it more important to understand the viral population in a sample as a whole rather than a single consensus sequence.

Evidence presented in the first section of the scenario was selected based on if it required a laboratory or scientific expertise to collect and analyze or not. Five types of evidence were included that would require such expertise to collect or analyze and five pieces of evidence that more traditional law enforcement or investigatory training would be required to collect and analyze were included. Environmental sampling was included based on the frequency at which interviewees mentioned the importance of such evidence in the prior aim. The delay in collecting these samples was added to mirror the latency in collecting environmental samples that often occurs due to delays in recognizing there is something spreading and where to look for samples. Postmortem samples were split into two groups, the original samples and the re-analysis, to see if participants judged the samples differently based on who did the analysis. Genomic and metagenomic evidence from other samples was included to assess how participants interacted with highly technical genetics data compared to more traditional types of data. The genetic analysis evidence was also included to compare how people interacted with the analysis of samples versus the samples from which the analysis was based (postmortem samples).

Part 2 of Scenario 1 was designed to make participants consider evidence specific to two suspects. One

suspect was written as having less direct access and expertise to genetically engineer a virus but more motive while the other suspect was written the opposite. Evidence for both suspects was made as close to equal as possible to avoid pointing guilt at one suspect more than the other. Facility inspections were included based on the importance of these inspections stated by participants in Aim 1. The thoroughness of these inspections was different for each facility to see if participants had different levels of trust based on the inspection depth. Because of the attention potential genetic engineering signatures had gotten during the COVID-19 origins discourse, signatures were included as evidence for this scenario.

Scenario 2

Where Scenario 1 was made to highly specific on one event, Scenario 2 was designed to be nebulous and represent a biological event that is long-term, not detected immediately, and of unknown origin. The same evidence was used for both parts of the survey for Scenario 2. Each section of the survey for Scenario 2 had a central question; in Part 1, participants were supposed to focus on determining whether they thought a naturally occurring, deliberate, or accidental event was being described. Part 2 asked participants to assess who or what may be responsible for the biological event occurring. The entirety of Scenario 2 is included in the appendix.

The scenario was not supposed to represent a clearly natural, deliberate, or accidental event. Dioxin was selected as the agent directly causing disease because it has been observed to cause health problems in humans before via diet. Other details such as the ASF vaccine plant and Nyx pesticide factory were included to offer possible explanations for an accidental contamination or deliberate actor. Details about geopolitics and misinformation were included to see if such details would impact participants analysis of the evidence. The same evidence was used for both parts of the scenario to see if interpretation of evidence changed based on asking what happened versus who did it.

4.2.2 Pilot Studies

The scenarios and two surveys were sent to volunteers to participate in a small pilot study. Results from the pilot study were used to determine if the scenarios were sufficiently understandable and to make sure they weren't too obviously pointing to one outcome. Based on feedback from pilot participants, the surveys were

updated.

4.2.3 Recruitment

Participants were invited to participate via email. Participants were chosen based on areas of expertise that would likely be picked to investigate a biological event. Ten participants who were interviewed in for the work completed in Chapter 1 were included as a participants in this work. Participants were offered a $200 honorarium for completing the survey.

4.2.4 Data Collection

A mixed-methods approach was used to design a survey to assess how participants interacted with the evidence presented in the scenarios. The questionnaire consisted of following types of questions: yes/no questions, ranking questions, multiple choice, and open-ended questions. The survey was designed to capture quantitative data through closed-ended questions and qualitative insights through open-ended questions. A copy of the survey can be found in the appendix.

The survey was created in Microsoft Forms. Participants were sent a link to the survey and PDFs of Scenario 1 and Scenario 2 via email after agreeing to participate. Participants were made aware that they could quit at any time and that their answers would be completely anonymous. Participants were given the option to voluntarily provide information on their area of expertise, nationality, years of experience, and gender. Informed consent was obtained from all participants, and their privacy and confidentiality were protected throughout the data collection and analysis process. Any identifying information was anonymized to ensure the confidentiality of participants' responses. This work was conducted under the purview of the Johns Hopkins Institutional Review Board as human subjects research (IRB number IRB00018728).

4.2.5 Survey Analysis

Three broad groups were made for comparison between areas of expertise, physical scientists, public health, and security/international relations/law (shortened to security going forward). Participants were given a list of areas of expertise and asked to select which areas they self-identified. Participants could select more than one area. Anyone selecting evolutionary and/or molecular biology, biotechnology, microbiology, or

environmental sciences was added to the physical scientists group. Anyone selecting public health or food safety was added to the public health group. Participants that selected international relations and/or diplomacy, forensics, biosecurity, chemical or nuclear security, or nonproliferation was added to the security group. Of the people who selected social science, they all picked at least one other area of expertise as well which was the basis of their groupings. All participants fit into at least one of these categories based on their self-identification. Some participants fit into two or all the categories.

R software was used for statistical analysis of quantitative data. For yes/no questions, responses were tabulated and the frequency of participants selecting each response was recorded. Descriptive statistics such as mean and standard deviation were calculated. Ranking questions were analyzed by assigning numerical values to each evidence option based on the rank provided by participants. The mean rank for each option was calculated to determine its relative importance among respondents. Thematic analysis was used on open-ended questions to identify recurring patterns, themes, and insights within the responses. Responses were coded and categorized based on common themes or topics using manual coding techniques.

4.3 Results

A pilot study was conducted followed by the main study. Results from the pilot studies conducted in June and July of 2022 informed the development of the main study. A total of 41 respondents participated in the main study between August 2022 and February 2023.

4.3.1 Pilot Study

After each scenario was developed, a pilot study was conducted using the scenario and a draft of the response survey. The purpose of the pilots was to determine if the scenarios and subsequent survey made sense to people who were not involved in their development and determine if the survey would generate the data needed to address the study's hypothesis. Twelve participants were included in the pilot for the first and second scenarios. Pilot participants were individuals with training and experience in public health, microbiology, communications, medicine, and/or the law. Participants were selected to mirror intended participants of the main study.

For the first scenario's pilot, nine participants stated Salle was killed with a biological weapon and three

stated Salle cause of not be attributed to the deliberate use of a biological weapon. When asked who was responsible for Salle's death, six of the twelve participants concluded that Roy Smith was responsible for throwing the dart, four participants selected Jim Little, and two participants selected the "Other" option. One participant who selected other stated in comments later in the survey that they didn't know who to pick and suggested adding an option choice to reflect that, which was done for the final survey.

In the survey for the second scenario, five participants stated they thought a deliberate event had occurred, four participants thought a natural event had occurred, and three participants said they did not know if it was deliberate or naturally occurring. In the second part of the scenario in an open ended question (rather than selecting from a list of options), six participants thought North Sol Cartel was responsible, three participants thought an the farmers responsible for taking care of animals were responsible for allowing the dioxin levels to be high, two participants said no organization or individual was responsible because it was entirely accidental, and one participant stated the vaccine manufacturing plant was at fault.

Because there was diversity in responses to both what occurred and who was responsible questions in the pilots for both scenarios, the authors determined that the scenarios were adequately nebulous for the purpose of this research.

Participants in both pilots reported the scenarios were clear and easy to follow. Many suggestions were given to improve the survey and clarify instructions for the exercise, which were then incorporated for the main study. Pilot participants were asked if they thought the included evidence was appropriate and all said they did find the evidence included in the scenarios appropriate.

Some participants in the pilot noted that they didn't know what the "correct" answers were and didn't have enough data given in the scenarios to make a confident conclusion about what happened. Along with the results from questions about what happened and who was responsible, the reported uncertainty was interpreted to mean that the scenarios were appropriate for the purposes of this study given that a scenario that was too obvious could have limited how people assessed each piece of evidence. Pilot study participants also reported that the evidence provided for each scenario was understandable and appropriate.

Following the pilot study, more questions were added to the survey to better assess how people were using the evidence in their analysis of the scenario. Pilot participants were not asked to provide a confidence level in their conclusion on a scale, but many tried to explain their confidence level in written responses. A

question asking participants to rank confidence was added for the main study based on the evident importance of relaying relative confidence in their assessments by the pilot participants.

Participants in the pilot frequently suggested they would feel more confident in making a decision if they "just had" one more piece of evidence. A question was then added for the main study asking what, if any, additional piece of evidence would have been helpful for their assessment.

Some pilot participants provided information about how they assessed credibility in their responses to other questions on the pilot survey while others did not touch on credibility and instead talked about their personal limitations in understanding the type of evidence or implications of a piece of evidence. Two areas of interest in this body of work are exploring how people assess credibility of evidence and determining how an individual's personal knowledge and understanding of a particular type of evidence contributes to how much weight that evidence holds in drawing a conclusion. To better address these areas, two additional questions were added to the main study, one asking for participants if there were any pieces of evidence that were difficult to assess or incorporate into their decision making and another asking how the participant assessed credibility of the evidence.

Following the pilot studies, the scenarios were not changed, and the survey instrument was updated before recruitment for the main study began.

4.3.2 Participant Demographics

Forty-one people participated in the main study between August 2022 and February 2023. Participants were asked to self-select their area(s) of expertise and were allowed to choose more than one area and write in other areas. The self-selected areas of expertise are reported in Table 4.1. Eight participants chose to write in other areas of expertise including evidence collection/preservation, demilitarization, conflict-zone investigations, cybersecurity, bioinformatics, plant pathology, agricultural sciences, emergency management, and pandemic preparedness, and emergency medicine. Participants were also asked to report their years of experience in their area(s) of expertise (Table 4.2).

Participants were given the option of reporting their gender identity and nationality. Twenty-one participants identified as a woman and 20 identified as a man. Nationalities included American, Argentinian, Australian, British, Bulgarian, Canadian, Chinese, Indian, Nigerian, Portuguese, Swiss, Rwandan, and Ugandan.

Area of Expertise	Number of Respondents
Evolutionary and/or Molecular Biology	6
Biotechnology	15
Microbiology (bacteriology, virology, etc.)	14
Social Science	8
Public Health	9
International Relations and/or Diplomacy	11
Law	0
Forensic Sciences	7
Biosecurity	28
Biosafety	13
Chemical or Nuclear Security	5
Nonproliferation	17
Food Safety	5
Environmental Sciences	4
Other	8

Table 4.1: Self-reported areas of expertise of participants in the evidence analysis exercise.

Years of Experience	Number of Respondents
0 to 5	4
6 to 10	13
11 to 20	13
21 to 30	8
31 to 40	2
41 to 50	1
Over 50	0

Table 4.2: Self-reported years of experience of participants in their area(s) of expertise.

4.3.3 Scenario 1

Scenario 1 was about the death of James Salle. The first set of questions asked participants to explore if evidence suggested Salle was deliberately killed with a biological weapon. The second set of questions asked participants to assume that he was killed with a bioweapon and determine who was responsible for his death.

Determining if there was an Assassination via Bioweapon

The participants were given ten different pieces of evidence to consider in determining if James Salle had been deliberately killed using a biological weapon. Five pieces of evidence would require a laboratory analysis and five pieces of evidence were more traditional types of evidence, such as witness statements, surveillance footage, or medical records. Each participant had to rank the ten pieces of evidence based on the influence of the evidence in their determination about whether Salle was killed with a bioweapon. No two pieces of evidence could receive the same rank. No two participants had identical rankings. The mean rank for each piece of evidence is reported in Table 4.3. A graphical representation of the distribution of ranks for each piece of evidence can be found in Figure 4.1.

Evidence Category	Evidence	Mean Rank	SD of Mean Rank	Aggregate Rank
Lab Derived	Metagenomic sequencing of swabs from dart	2.15	1.50	1
	Genomic analysis of post-mortem samples	2.22	1.50	2
	Postmortem samples collected from Xenon Medical Examiner	3.88	1.21	3
	Postmortem samples reanalysis	4.00	1.70	4
	Environmental sampling of venue	7.32	1.94	7
Non-lab Derived	Medical records from the hospital where Salle died	4.39	2.09	5
	Interviews with Salle's personal staff and close contacts	6.90	2.29	6
	Witness statements	7.39	1.56	8
	Event venue surveillance video recording	7.51	1.53	9
	Surveillance footage from Argon Capital City CCTV	8.66	1.44	10

Table 4.3: Mean and aggregate rank of different types of evidence's influence on determining if Salle was assassinated with a bioweapon.

A Shapiro- Wilk Test was done to determine if the ranking data fit a normal distribution. The test statistic was 0.98 and the p value was 0.00023, so the null hypothesis that the data fit a normal distribution was rejected. Because the data did not follow a normal distribution, a Kruskal-Wallis Test was used to determine if there was a significant difference in mean ranks of the ten pieces of evidence. The test statistic

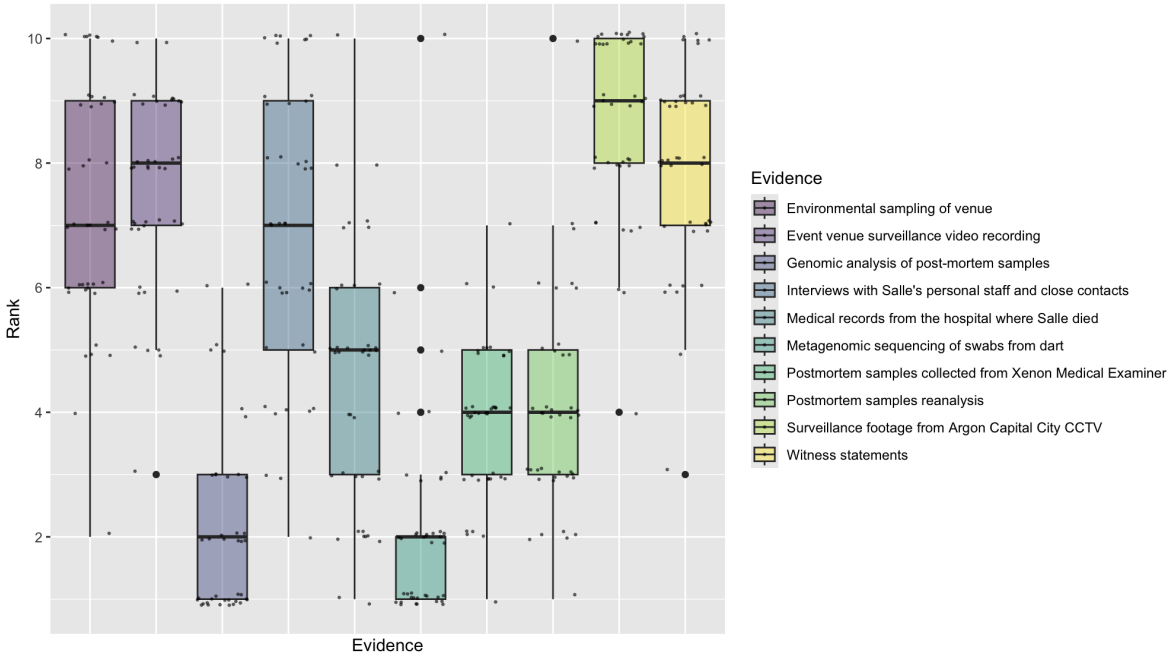

Figure 4.1: Distribution of rankings of evidence relative influence on determining if Salle was deliberately killed with a bioweapon.

H was calculated to be 264.5 with 9 degrees of freedom and a p value of less than $2.2\text{x}10^{-16}$, so the null hypothesis that the rankings for each piece of evidence originated from the same distribution was rejected. We conclude that at least one type of evidence has a median ranking different from the other pieces of evidence.

A Wilcoxon Sign Ranked Test with continuity correction was used to determine which pieces of evidence had significant differences in their median rankings. The pairings of evidence with significantly different median ranks were determined based on which pairings had a test statistic W less than the p value of 0.005. Test statistics for each pairing are reported in the appendix (Table A1).

Of the 41 participants, a total of 34 said there was enough evidence to conclude Salle died from a deliberate attack using a biological weapon. Five of the remaining seven participants stated that there was a lot of evidence pointing this conclusion, but they could not definitively say that Salle's death was a deliberate act because not all other potential causes of his death had been ruled out. The final two responses did not indicate they thought there was any evidence pointing to this conclusion.

The percentage of respondents identifying as experts in each of the three overarching areas (physical

sciences, public health, and security) who responded that there was enough evidence to conclude Salle was deliberately killed with a bioweapon was compared to the percentage of participants not identifying as an expert in the given area. For physical scientists, 81% of physical scientists said 'yes' while 82.4% of non-physical scientists said 'yes'. For public health, 84% of public health experts said 'yes' and 78% of non-public health experts said 'yes'. In the security group, 90% of these experts said 'yes' and 81% of people not identifying as a security, international relations, or law expert said 'yes'. For each of these comparisons, a chi-square test was done to test for significant differences between those identifying as an expert in the area versus those who did not identify as an expert in the area. No area had a significant difference in the percentage of 'yes' responses between expert versus non-expert (Table 4.4).

Area of Expertise	**Experts Saying 'Yes'** (%)	**Non-Experts Saying 'Yes'** (%)	**Chi-Square Statistic**	**P Value**
Physical Sciences	82	84	2.01×10^{-30}	1
Public Health	84	78	1.29×10^{-31}	1
Security	90	81	0.04	0.84

Table 4.4: Comparison of experts vs. non-experts in each major area of expertise responding 'yes' to whether there is enough evidence to conclude Salle was deliberately killed using a bioweapon.

Participants were asked to rank on a scale of zero to ten how confident they felt in concluding if Salle's death was a deliberate assassination with a bioweapon. Answers ranged from two to ten with an average of 8.2 (standard deviation of 1.6). There were only two scores below a six; a two and a five, which were from the same two people who did not express there was evidence pointing towards concluding Salle was deliberately killed with a bioweapon in a previous question. The individual that ranked a two on this question identified as an expert in international relations and nonproliferation. The participant that ranked a five on this question identified as an expert in biosecurity and biosafety. The distribution in scores is shown in Figure 4.2.

When asked to expand on why they thought Salle was assassinated or not, the most common reason stated for believing it was a biological attack were the genetic analyses of the rabies virus from the dart and postmortem samples, which is reflected in the ranking data. The degree of similarity between rabies sequences from the dart and the postmortem sample were "past the point of coincidence" for many participants. Fifteen participants pointed out that multiple copies of a pathogenicity glycoprotein gene in addition to multiple mutations known to increase pathogenicity in both copies were enough to confirm someone had

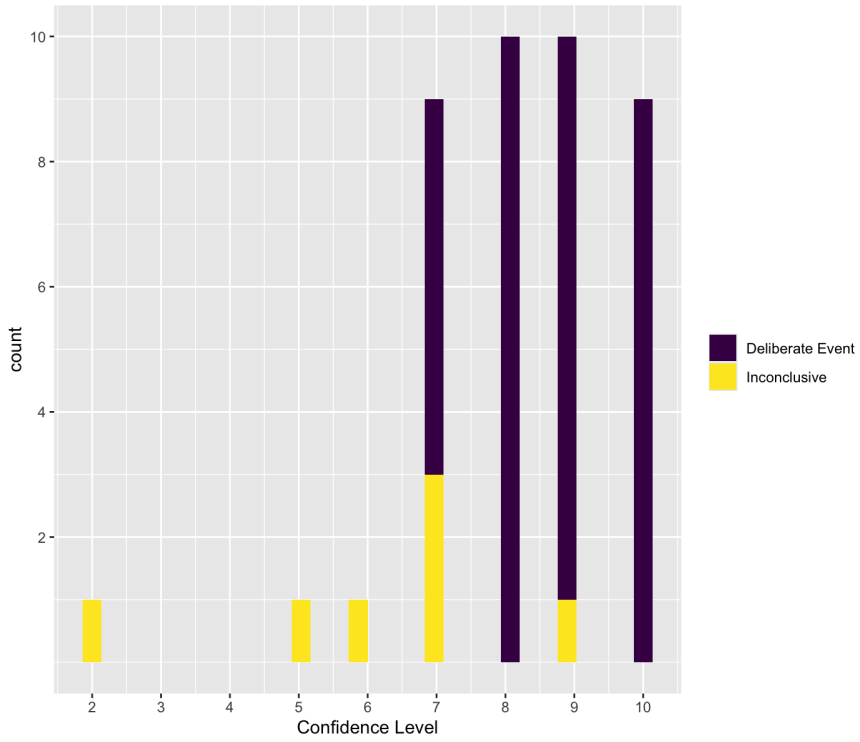

Figure 4.2: Distribution of confidence level of participants when determining how confident they were that James Salle had been assassinated with a bioweapon.

genetically engineered part of the virus.

Of respondents that were not confident in concluding Salle had deliberately been infected, they universally mentioned that there were too many other pathways for infection that had not been ruled out. Reliability and accuracy of Salle's contacts' interviews was questioned frequently. Medical records from before the event was also requested five times to determine if there were any prior signs of disease pre-dating the event where a dart was thrown at him. One participant who did not think there was enough information to conclude a deliberate event stated the biggest reason they didn't think it was deliberate was because there was no way a rabies virus would be environmentally robust enough to survive on a dry dart at an outdoor venue long enough.

Of the two participants that had confidence less than 6, one gave a low score because they felt there was not enough evidence to eliminate other potential routes of rabies exposure for Salle, such as if he had an exotic wildlife trafficking hobby. However, the same participant said the evidence provided in the scenario:

> "point(s) towards an Occam's Razor-esque conclusion: given that a dart which was found to contain fragments of rabies nucleic acid appeared was used in an attack on James Salle, who then ended up dying from rabies...there seems to be a clear connection between the dart, the victim, and the manner of death. This would indicate that this COULD be an example of the use of a biological weapon. However, the evidence provided in this scenario alone does not provide enough information to eliminate other potential situations where Salle may have been exposed to rabies in another fashion."

The most cited evidence supporting the conclusion that this was an assassination with a bioweapon was the genetic engineering signatures in the rabies genome. Many participants assumed that the pathogenicity mutations were evidence of genetic engineering. However, these mutations could arise naturally for a virus. Only one participant pointed out that the mutations could be naturally acquired by the virus. This participant, who self-identified as a microbiologist, was also the only one to discuss the possibility of a virus population being able to sustain double copies of the glycoprotein gene within their genome. While many participants were using possible evidence of genetic engineering as proof of intent, this participant was very careful to do the opposite in their analysis. They stated that these changes in the genetic of rabies virus could be indicative of unique selection pressures on the virus population that had not been observed before, which

could mean there are changes in the expected natural transmission of the virus that could account for Salle's infection. This individual gave a six when asked for their confidence that this was a deliberate event.

Twelve participants seemed to focus on the potential viral engineering entirely as evidence of deliberate based on their written explanations. These participants ranked the genomic analysis of the samples (which included the information about increased pathogenicity mutations and double copy of the glycoprotein gene). Of these twelve participants, they ranked the genomic and metagenomic evidence options in first and second place or vice versa, followed by the five more traditional evidence types, then the postmortem samples and environmental samples. Of the other 22 participants who were confident a deliberate event had occurred, all of them ranked at least three of the five lab-based evidence samples in their top five for the ranking. Many of these 22 participants focused on the fact that there was no evidence of a rabies bite from an animal (according to interviews with Salle's contacts) in their written responses and that the incubation time for rabies fit for the event described. Four of these 22 participants stated that the metagenomic analysis of the dart was worthless as evidence because there was no way to tell if Salle was infected by the dart or if the virus got on the dart from an already infected Salle.

Eight participants noted the environmental samples were irrelevant because too much time had passed between the event and sample collection (43 days according to the scenario); many of these individuals ranked the environmental samples in the bottom 3 in their rankings. Two participants mentioned the possibility of evidence falsification, but both also felt it was unlikely given that samples taken from two different sources (dart and autopsy) had similar results.

Credibility of Evidence

In addition to asking participants about the overall impact the evidence in the scenario had on their determinations, they were also asked about credibility of the evidence. The overall attitude of the responses can best be summed up by this quote, "witness statements and statements by personal contacts are notoriously unreliable. Very little credibility. Surveillance footage is credible but not particularly useful. The genomic and meta-genomic analysis (and re-analysis) is most credible and trustworthy." Thirty-six participants mentioned these hierarchical levels of evidence credibility in their responses.

Ten participants said credibility was high because samples had been collected and/or analyzed by mul-

tiple teams. One participant highlighted this by saying, “many pieces of evidence were corroborated with independently and repeatedly verifiable documentation and testing results.” Another participant emphasized the agreement between analyses as being critical for credibility; “ knowing that the results of many of the analyses agree with each other, I am assigning higher credibility than if they disagreed.” Three participants mentioned they would consider the credibility of the evidence to be high overall if the chain of custody of the samples could be validated.

Fifteen participants said the genomic data was credible and convincing over all other evidence Five participants stated that the genomic data was either not credible or not convincing due to inability to determine when the rabies RNA got on the dart and the low sequence similarity. While one participant called the sequences on the dart a “smoking gun”, another called it a “red herring.” Table 4.5 includes exemplary quotes participants wrote in response to how credible they found the evidence.

“Genomic data seems like it has the potential to provide concrete evidence, but I don’t know enough about it to understand what I’m reading. For example, I don’t know how much of a difference there is between 83% and 99% in terms of the degree of confidence that specimens match. Or how much it matters that he dart specimens were 99.2% similar to the consensus sequence while the two analyses of the Salle specimens were only 98.7% similar. Is the 86-97% similarity between the University Labs rabies specimens and the Salle specimens conclusive enough? How much better is that than the 83% in the IOM lab? If you took a random natural specimen of rabies, what would the similarity be? What is this acceptable range? I have no idea.”
"Where I find challenges in terms of associating specific suspects with the attack is based on the variability in the sequencing data. While I am not familiar with the intra-patient stability of the rabies virus, it seems odd that only 90 of the genome’s sequences had 2 G-protein copies. Assuming that we were starting from a clonal rabies virus and this mutation helps infectivity, why the drop off in % ? Similarly, why the difference in the results between the post mortem analysis (98.7%) and metagenomic sequencing (99.2%)? Is this perhaps an issue with the diagnostic NGS process? For the 4 SNPs, could we be sure that they were all from a single virus based on the reads, or is it possible there was a mix of different viruses in the individual? A thorough analysis of the sequencing approaches and analysis I think would need to be done to avoid potential analysis artifacts.”
“Sequence variabilities are a little higher than I might expect.”
“After sequencing, samples from the IOM rabies lab grouped separately from those taken from Salle’s body, and their sequence similarity was only 83%), suggesting that they were not from the same source.
“The metagenomic sequencing of swabs from the dart which showed 99.2% nucleotide sequence similarity with the consensus sequence created from the autopsy samples. The genomic evidence is quite convincing and the research seems very thorough.”
“We had the sequence from rabies, done almost simultaneously by different entities. One is the Advisory Group created after the event, by subject matter experts, and then the lab where he was hospitalized. Having a coincidence of almost 99% between both entities, makes the evidence to be credible.”

Table 4.5: Exemplary responses from participants concerning how they interpreted the crediblity of the evidence in the scenario.

Additional Evidence

When asked what single piece of additional evidence they would like in order to determine if a deliberate event occurred, 29 participants said they wanted more details about the dart, such as its structure, how it would carry a payload, and the manufacturer of the dart. Four participants asked for metagenomic analysis of samples from Salle to compare to the metagenomic analysis taken from the dart. One additional person asked for a quantification of human to viral RNA found on the sample. Confirmation that infectious rabies was on the dart, whether or not the metagenetic analysis from the dart showed two copies of the glycoprotein gene, and information on the likelihood of how likely rabies virus is to be found on a dart that punctures skin, if the person punctured already has asymptomatic rabies were each requested by a participant. Three participants asked for more extensive medical records from Salle.

Determining Who was Responsible for the Assassination

For section 2 of Scenario 1, participants were instructed to assume a deliberate event had occurred and focus on who may be responsible for the assassination. Participants were first asked to select who they thought threw the dart based on the next set of evidence. Approximately 22% of participants picked Jim Little, 37% picked Roy Smith, 37% said another person, one person said no one threw a dart and one person didn't answer the question (Tables 4.6 and 4.7). Interestingly, when broken down by areas of expertise, 63.1% participants identifying as scientists chose an individual suspect whereas only 54% of non-scientists chose a suspect. The overall group average for choosing a specific suspect was 59%. The percentage of public health to non-public health experts choosing a named suspect was 55% to 59.5%, respectively. The percentage of public health to non-public health experts choosing a specific suspect was 58.1% to 60%, respectively.

Suspects	Number of Participants	Overall Percentage
Jim Little	9	22%
Roy Smith	15	37%
Another person	15	37%
No one threw a dart	1	2.4%

Table 4.6: Participant choices of who threw the dart at Salle.

Certainty in choice of suspect was low. When describing their confidence in selecting who threw the dart, 32 participants said there wasn't enough information to feel confident in any choice, as reflected when

Suspects	Scientists/Non-Scientists	Public Health/Non-Public Health	Security/Non-Security
Jim Little	36.8% / 9%	11.1% / 25%	12.9% / 50%
Roy Smith	26.3% / 45.5%	44.4% / 34.4%	45.2% / 10%
Another person	31.6% / 40.9%	44.4% / 34.4%	35.5% / 40%
No one threw a dart	5.4% / 0%	0% / 3.1%	3.2% / 0%

Table 4.7: Difference in percentage of people choosing a given suspect by self identified areas of expertise

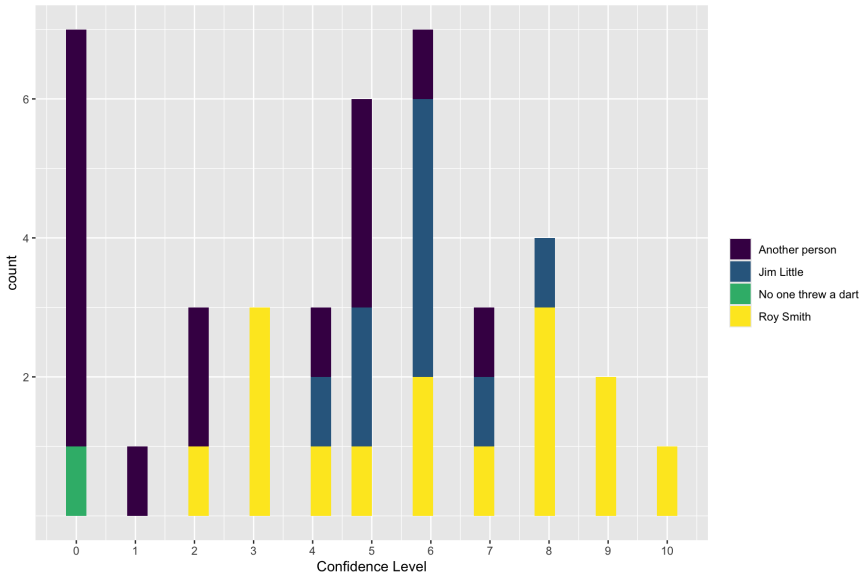

Figure 4.3: Distribution of confidence level of participants when determining how confident they were in choosing a suspect for who threw the dart.

asked to describe their confidence on a scale of zero to ten (Figure 4.3). The average confidence level was 4.44 (standard deviation of 2.95) and the range of responses was zero to ten.

One of these 32 chose not to answer the question, one said no one threw a dart because they didn't have an option to say, "I don't know," and fourteen said another person threw the dart. Of the participants who chose another person as the dart thrower, three said they chose that option because statistically unlikely that it was either Smith or Little who threw the dart out of the pool of 68 people who were not filmed at the event. Twelve people said there wasn't enough evidence to choose Little or Smith, so they chose another person. Seven participants also stated that while Little or Smith may have been involved with the assassination plot, that did not mean they threw the dart themselves. Highlighting the potential difference between being involved and throwing the dart, one participant stated, "there is little evidence connecting these two persons to the dart and very little evidence connecting them to the act of projecting the dart into Salle." Another

statement about Smith said, “if Smith was involved, it seems that he is plausibly not the engineer, but rather knew the engineer.”

Overwhelmingly, participants focused on a suspects’ access to rabies viruses and the genetic comparisons of that virus to the samples taken from Salle and the dart when considering suspects. Thirty-five respondents mentioned access or similarity of viral RNA when describing how they came to their conclusion on who threw the dart and of those 35, 25 stated the similarity between strains was their primary motivation for either choosing Jim Little or Roy Smith or specifically not choosing one of them.

Six participants mentioned motivations of the suspects in their reasoning with two participants saying they would have chosen Little but there wasn’t evidence of a motive for him beyond disparaging remarks on social media, unlike Smith. When discussing Smith’s possible motives, five participants noted that Smith was belligerent in his interviews and highly vocal in his opinion about Salle’s pipeline and the potential impacts on the environment. One participant said, “Roy [Smith] probably sees death from a disease associated with wild animals as some sort of pyric justice... Nature getting back at Salle... or something.”

Participants were again asked to rank ten pieces of new evidence based on how influential the evidence was in their determination of who threw the dart. No two pieces of evidence could receive the same rank. No two participants had identical rankings. The mean rank for each piece of evidence is reported in Table 4.8 and the distributions of rankings are illustrated in Figure 4.4.

Evidence	**Mean Rank**	**Standard Deviation of Mean Rank**
Genetic Engineering Signatures in the Rabies Lab Samples	3.23	2.23
University Laboratory Inspection	3.54	2.44
Samples Collected from the IOM Rabies Lab	3.9	1.98
Samples Collected from Little’s Lab	4.62	2.11
IOM keycard logs and security footage	4.62	3.01
Biotech Facility Inspection	6.59	1.79
Rabies epidemiology records	6.67	2.74
Interview with Jim Little	6.72	2.89
Interview with Roy Smith	6.79	2.42
Social Media Analysis	8.33	1.96

Table 4.8: Mean of different pieces of evidence’s influence on determining who was responsible for throwing the dart that killed Salle.

A Shapiro- Wilk Test was done to determine if the ranking data fit a normal distribution. The test

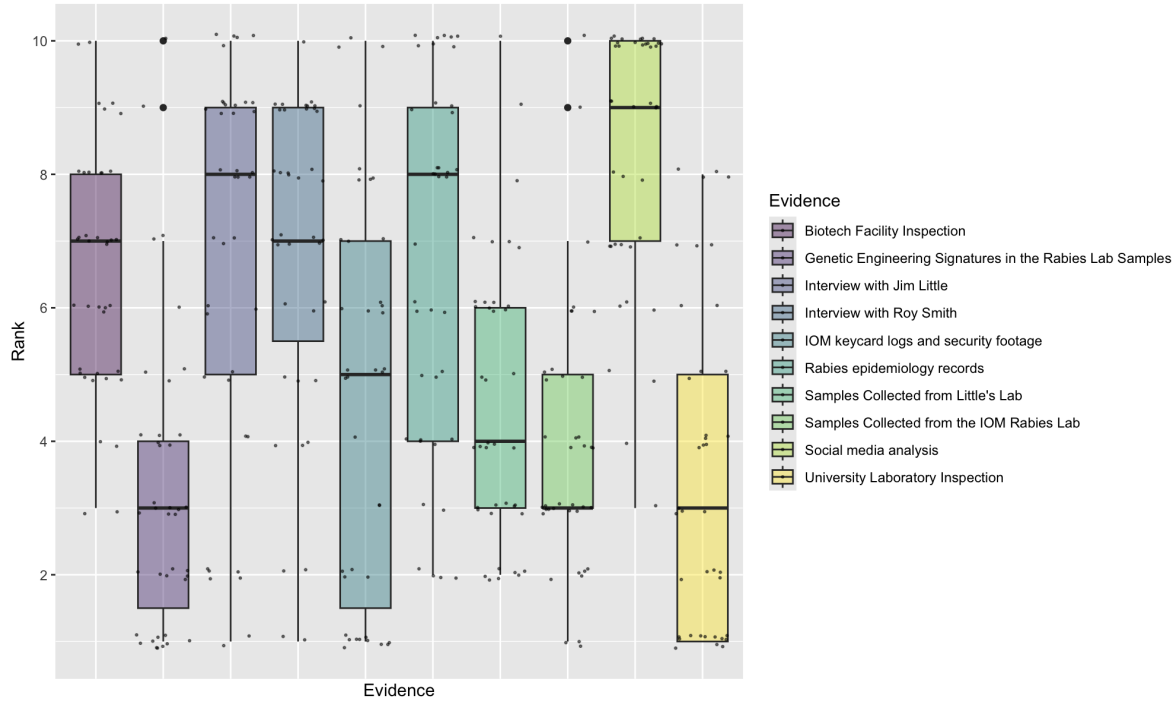

Figure 4.4: Distribution of rankings of evidence relative influence on determining who threw the dart at Salle.

statistic was 0.99 and the p value was 0.01058, so the null hypothesis that the data fit a normal distribution was rejected. Therefore, a Kruskal-Wallis Test was used to determine if there was a significant difference in mean ranks of the ten pieces of evidence. The test statistic H was calculated to be 126.68 with 9 degrees of freedom and a p value of less than $2.2\text{x}10^{-16}$, so the null hypothesis that the rankings for each piece of evidence originated from the same distribution was rejected. Therefore, at least one piece of evidence has a median ranking different from the other pieces of evidence.

A Wilcoxon Sign Ranked Test with continuity correction was used to determine which pieces of evidence had significant differences in their median rankings. The pairings of evidence with significantly different median ranks were determined based on which pairings had a test statistic W less than the p value of 0.005. Test statistics for each pairing are reported in the appendix (Table A2).

Participants identifying as scientists all had either one of the laboratory inspections, the lab keycard access logs and security footage, or analysis of one of the samples as their top ranked evidence. For people not identifying as a scientist, if their top rank was not one of the five listed above, then it was an interview with one of the suspects.

When describing what evidence, if any, were exonerating for any suspect, seventeen participants said there was no evidence exonerating either suspect. Twelve of these participants also felt there was weak or no evidence pointing towards either named suspect. Two participants didn't answer the question. Of the remaining 22 participants, four participants said the genetic difference between samples and different number of glycoprotein genes together was exonerating of Little. One participant said the genetic difference between samples and different number of glycoprotein genes together exonerated both Smith and Little. Three participants said the genetic distance between samples from Salle and the IOM lab were exonerating of Little while one participant said the genetic distance between Salle's samples and the university laboratory's publications exonerated Smith. Two people said the lack of a double glycoprotein gene in the IOM lab samples exonerated Little, one person said the lack of a double glycoprotein gene in the biotech facility stocks exonerated Smith, and two people said the lack of the double glycoprotein gene in both samples exonerated both men.

One person said the epidemiological data concerning rabies in the area exonerated both men because the records showed it was possible that Salle already had rabies and the viral genome got on the dart when it pierced his skin.

Three participants said facility inspections exonerated Smith. However, three participants specifically stated that the facility inspections were not adequate for drawing any conclusions, especially the university laboratory inspection. The reliance on an online database for sequences to compare to those from Salle was highlighted as a serious flaw in the investigation by one participant. Five participants mentioned that there needed to be more searching for alternative work sites for both men; personal DIYbio labs were mentioned as possible sites where either man could have genetically modified the virus. When ranking top evidence for making their decision on attribution, twelve participants ranked the university laboratory inspection as their most influential choice.

When discussing the genetic analysis from samples collected following facility inspections, four participants said samples needed to be collected from the university laboratory. There were also several comments about the genetic distance between viral sequences from the university laboratory and Salle. Two participants noted that 96.9% homology is impossible from chance. In describing their thought process in determining fault, one person wrote,

> "the number of degrees of freedom of possible pathogenesis enhancing mutations is HUGE. That one suspect is associated with all four of those same mutations (through a friend) is statistically impossible, add in that the friend has the relevant expertise, access to a lab, and would know how to clean that lab to make sure it sampled negative later, and it's a much clearer case of MEANS."

Three additional participants said it would be too difficult to engineer a rabies virus with considerable divergence from a baseline strain, therefore they thought the strain used to kill Salle must have come from the university laboratory.

Keycard access to the IOM lab was ranked as the number one piece of evidence by ten people, seven of which identified as scientists, two as a security-related area expert, and one as a public health expert. Fifteen people mentioned the keycard access when considering if Little or an associate of his could have been responsible for generating the virus on the dart. Others were dismissive of the keycard data. One participant said the visit frequency and duration from the key card data was not a long enough time to do viral engineering, so either he was picking up a sample someone else produced or he was just visiting and chatting with friends. Another participant said,

> "Were there security irregularities? Yes... there are ALWAYS some irregularities. That little tried to enter a door that he wasn't allowed to open... Who cares? They wouldn't have these doors locked if researchers could be expected to keep track of where they are and aren't allowed to go. Similarly the unidentified bag... who can account for every article they carried just YESTERDAY?"

Evidence Credibility

Attitudes towards credibility of evidence in this section differed among participants. Of the participants that answered the question, nineteen found the evidence credible in large part due to the IOM keycard logs and security footage, inspections of facilities, and nature of genetic analysis ("genetics doesn't lie") and most excluded the interviews from their conclusion of overall credibility. Eight of these nineteen said that while the evidence was credible, it was not conclusive or comprehensive. Ten respondents said the evidence was partially or mildly credible and two said the evidence was not at all credible. For the two individuals stating

the evidence was not credible, both pointed out that there was more subjective evidence included in this section of Scenario 1 than the previous section. Both also said the overall investigation, but especially the inspections, were not thorough enough to be considered credible.

According to one participant who stated they were an expert in law enforcement policy, it was challenging to assess the technical evidence and therefore they were unable to comment on its credibility. Two respondents mentioned they would feel more confident in evidence credibility if there was more information about chain of custody and standard procedures used by investigators included in the scenarios. Interestingly, one respondent mentioned the unique situation of biological events in their response,

> "Contrary to other criminal investigations, crimes using biological agents do not comply with Standard Operational Protocols (SOP). The investigations must be conducted based on previous expertise and reported cases and depend highly on the agent and the conditions. The credibility assessment was based on my knowledge of previous criminal cases and focused on the analysis performed and on the timely application of such techniques."

Additional Evidence

When asked what single additional piece of evidence participants would want to see, all had at least one request. The most common request (from eight participants) was additional searches of property associated with the suspects including their homes, workplaces, and any laboratory space for which they or their contacts had access. Five participants wanted more in-depth interviews and investigations into contacts of suspects. Further analysis of the dart and fingerprints from the dart each had four requests. Alibies or location logs for the suspects, investigations into rabies virus natural sequence space and mutability, and more CCTV were requested by three participants each. Sequences of naturally occurring rabies cases in the area, information on the suspects' dart throwing ability, and searches for the missing badge and a dark hoodie were requested two times each. Order records from the facilities related to rabies primers, an admission of guilt, testing for preventative vaccination of suspects, and background checks on everyone were each requested by one participant.

Overall Comments on Scenario 1

The survey asked participants to describe any evidence they found inappropriate or difficult to analyze in either section of the scenario. Eight people said the genetic analysis components were difficult for them to incorporate into their thinking for their respective sections. Eight people also said the environmental sampling data was inappropriate because it was taken too late after the event occurred. Seven people said the social media analysis was not difficult to use for their analysis because it was hard to verify. The most common concern stated throughout the survey with this scenario was that the pool of suspects was narrowed down too quickly to Little and Smith and that the narrowing to these two individuals was problematic because it heavily relied on weak evidence (the social media analysis).

Generally, participants felt it was possible to identify with confidence the agent causing Salle's death but not the origin of the agent. One participant stated that because there was strong evidence that Salle had rabies when he died, they were confident in that part of the investigation, but all other conclusions were questionable because, "the forensic analysis and the criminal investigation process would need to be integrated and not consecutive or separated from each other."

Three participants said their ability to consider the questions asked in the survey would be strengthened by inclusion of background check information on suspects and their contacts, information of the political climate in the region, and intelligence reporting from the region.

Thirty-five participants stated that they prioritized genetic analyses over all other types of evidence when assessing credibility, even if this evidence wasn't the most influential in their decision making, because genetic evidence was entirely objective.

When discussing how the scenario could reflect a real-life investigation, seven participants mentioned the importance of "which labs that undertook the sequencing and having trust in the analysis locations would probably be important in real world."

4.3.4 Scenario 2

Scenario 2 revolves around the countries of Sol, Nyx, and Hemera and unusual disease trends in animal and human populations in these countries. The cause of the increased morbidity and mortality is eventually believed to be associated with dioxin exposures. The origin of the dioxin in the food supply is unknown

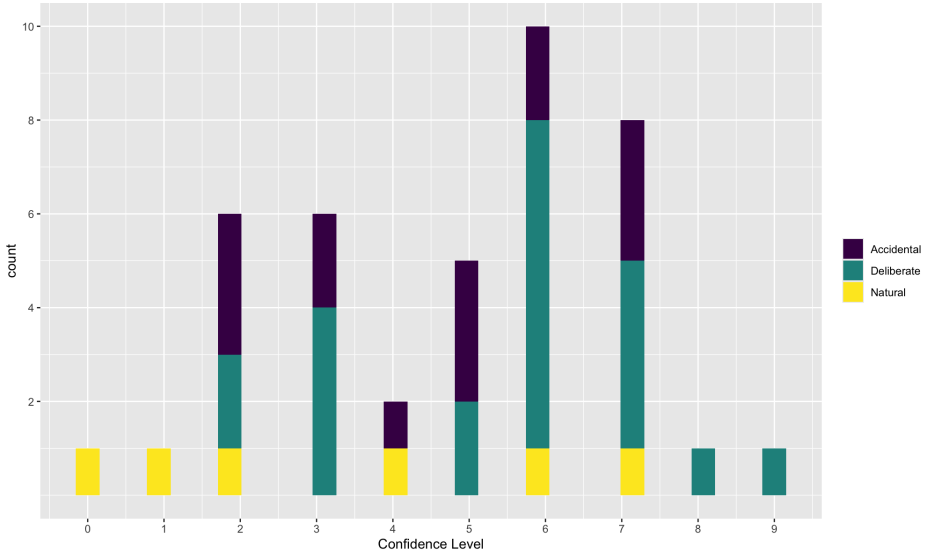

Figure 4.5: Distribution of confidence levels of participants when determining how confident they were in determining the nature of the event described in Scenario 2.

but rumors of the cause being due to a bioweapon spread and lead to a multilateral investigation. This scenario included ten pieces of evidence and participants were first asked to use the evidence to assess if a man-influenced event had occurred followed by a second part looking at possible attribution.

Determining the Nature of the Event

The first questions participants were asked for this scenario was if they thought the scenario described an event that was a man-made event or naturally occurring. Thirty-five participants selected, "at least some of the cases of disease are associated with a deliberate or accidental event." The remaining six participants selected the option stating all cases of illness were due to naturally occurring events. Of the 35 people believing there was an anthropogenic cause for the observed illnesses, fourteen suspected it was an accidental event and 21 suspected it was a deliberate event. On a scale of zero to ten, the average confidence level in their conclusion was 4.8 (standard deviation 2.16) with a range from zero to nine (Figure 4.5).

Naturally occurring verses man-made

Participants were asked to describe their reasoning behind picking each type of event. Of the six participants who thought the scenario described a naturally occurring event, one person said they chose naturally occur-

ring because the environment described in the scenario sounded like the "environment of middle-income countries where the type of development led to things like incinerating waste or industries without enough contamination control." Another respondent suggested that it was a natural event that could be explained by differences in wild verses farmed pigs' genetics, where pigs raised on a farm are more susceptible to something. The other four participants said there wasn't enough data to suggest anything other than a naturally occurring event.

The most cited reason for not believing the event was naturally occurring was the disparity in the disease rates between farm-raised pigs and wild pigs with 20 participants. A lack of elevated dioxin levels in the environment was the next most mentioned reason for picking an anthropogenic cause with eleven participants. A lack of a natural source of dioxin and the apparent association between the ASF vaccine and cases in pigs were both mentioned four times in the reasoning. Intelligence suggesting abnormal behavior from the North Sol Cartel, the abnormal disease clusters, and severity of disease in humans were each stated three times. Other reasons for determining the event were man-made include the temporal relationship of increases in dioxin-related consequences, transcriptomic markers of dioxin exposure in humans, interviews with the local vets and intelligence signals suggesting suspicious activity, and the initial unwillingness of Nyx to participate in a multilateral investigation. Additionally, one participant stated that because three different countries were each facing the same problem, it could not be a naturally occurring event.

Accidental versus Deliberate

Of the respondents who did believe there was evidence that some of the cases could be caused a man-made event, six of them said it was a "toss up" between deliberate and accidental but without more evidence to suggest either way, they would default with accidental. Another two participants said that without explicit evidence of a deliberate event, they would not select deliberate. A lack of apparent targeting of disease among humans was cited twice. Lack of acute toxicity with proximity to a suspected origin of the event was another reason for concluding accidental. Three people suggested there may be accidental feed contamination with one person even referencing a time when dioxin levels were increased in pigs in Ireland due to contaminated feed. One respondent clearly laid out their deliberations as follows,

"The general elevation of new cases across Year 1 and 2 across all areas indicates actual

cause has wide geographic spread and lacks targeting. The year 3 dop off for Sol is inconclusive given the focus on stringent health practices identified in some of the evidence. Additionally, the river flow would likely remove a one-off environmental contaminant from the Nyx region first, potentially contributing to the early drop off rate across year 3-4. However, there is low confidence in this data as it was reported by Nyx to WHO. "

One respondent stated they would not be able to make a conclusion about accidental verses deliberate without "video footage of someone purposefully contaminating a pesticide supply or the like."

The most frequently cited reasoning for determining an event was deliberate was because of the behavior of the North Sol Cartel, such as the precautions they took to keep their pigs separate and only eat their pigs, which was listed in the reasoning of fourteen participants. Four participants also stated that in addition to their behavior, the cartel had clear motive, like to expand influence, or undermine the Sol government, which further suggested a deliberate event. The disparity in dioxin levels in wild compared to farm raised pigs was cited three times as evidence of a deliberate rather than accidental event.

Four participants said the difference in impact was crucial in their determination with one saying, "weak but suggestive evidence that the North Sol Cartel was not affected," implying that the disproportionate amount of impact was itself suggestive of a deliberate event. Another participant suggested that because dioxin linked genes were upregulated and transcription profiles altered in humans, it was a deliberate event. Three participants noted that there seemed to be an association between the ASF vaccine and pathogenicity, but doubted a vaccine could cause as widespread cases as reported in the scenario without an entire batch of vaccine being affected, and an entire batch wouldn't be affected without deliberate action. The lack of environmental exposure was cited as evidence of a deliberate event as well.

Participants were given ten different pieces of evidence to consider when determining which type of biological event was occurring. Each participant had to rank all ten pieces of evidence based on how influential the evidence was in their determination of event type. No two pieces of evidence could receive the same rank. No two participants had identical rankings. The mean rank for each piece of evidence is reported in Table 4.9. A graphical representation of the distribution of ranks for each piece of evidence can be found in Figure 4.6.

A Shapiro- Wilk Test was done to determine if the ranking data fit a normal distribution. The test statistic

Evidence	Mean Rank	Standard Deviation of Mean Rank
Surveying pigs for dioxin bioaccumulation	3.45	2.42
Environmental sampling for dioxins	4.85	2.55
Transcriptomic analysis of samples from current patients	5.00	2.68
Medical records from patients in Hemera, Nyx, and Sol	5.28	2.61
Satellite images of Sol's cartels' territories	5.30	2.73
Intelligence Signals in the Region	5.58	2.88
Epidemiological data from national authorities	5.92	2.68
Interviews with local veterinarians	5.92	2.68
Satellite imagery analysis of chemical processing plants and other facilities in Nyx	6.92	2.60
Previous published research article	7.05	3.14

Table 4.9: Mean rank of different pieces of evidence's influence on determining the type of biological event occurring in Scenario 2.

was 0.98 and the p value was 3.72×10^{-6}, so the null hypothesis that the data fit a normal distribution was rejected. Because the data did not follow a normal distribution, a Kruskal-Wallis Test was used to determine if there was a significant difference in mean ranks of the ten pieces of evidence. The test statistic H was calculated to be 46.5 with 9 degrees of freedom and a p value of less than 4.94×10^{-7}, so the null hypothesis that the rankings for each piece of evidence originated from the same distribution was rejected. At least one type of evidence has a median ranking different from the other pieces of evidence.

A Wilcoxon Sign Ranked Test with continuity correction was used to determine which pieces of evidence had significant differences in their median rankings. The pairings of evidence with significantly different median ranks were determined based on which pairings had a test statistic W less than the p value of 0.005. Test statistics for each pairing are reported in the appendix (Table A3). Only three pairings of evidence were found to have significantly different mean ranks: the pig survey data compared to the satellite images, the pig survey data compared to the previous research, and pig survey compared to interviews with vets.

Unsurprisingly given the prior explanations of how participants determined the type of event, the pig surveying data had the highest mean rank and was the most frequently picked top rank. However, there was not a significant difference in rank between any other types of evidence. There was clear overlap in the second and third quartiles for all evidence comparisons other than the pig survey data compared to the previous research.

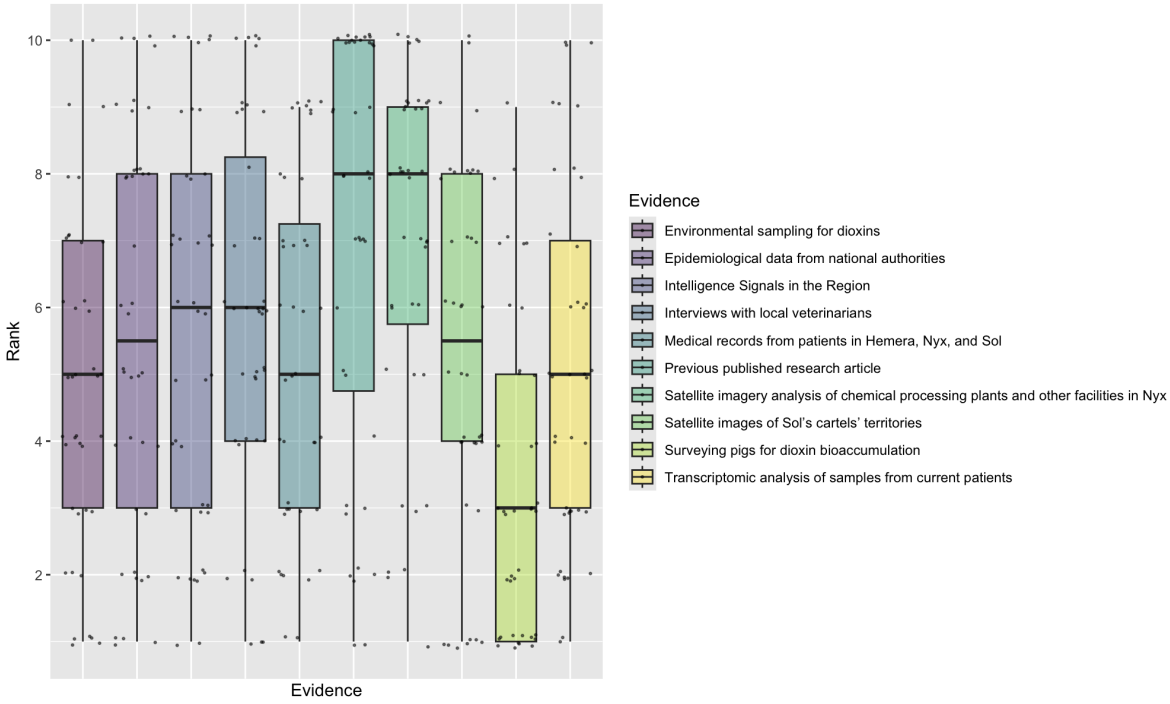

Figure 4.6: Graphical representation of the distribution of rankings of evidence relative influence on determining the type of biological event described in Scenario 2

Additional Evidence

Participants were asked to suggest one additional piece of evidence that would have helped them determine what type of biological event may have occurred in Scenario 2. Four participants did not respond to the question. Seven participants requested an investigation into the ASF vaccine manufacturing facility, its processes, and the staff. Six people requested testing of the ASF vaccine itself. Six people requested pig feed across the countries be analyzed, with one participant providing citations for past occurrences of feed contamination leading to adverse human health outcomes (171). Six participants requested more information from intelligence sources, including requests for intelligence about the cartels, geopolitical landscape of the region, trade (particularly pork trade), and even weather and climate trends in the area.

Four participants requested more sampling of pigs in the different areas of Sol, including the territory that is government controlled. Two additional participants requested genetic testing of the pigs in Sol with one specifying they wanted the testing done pre- and post-ASF vaccination of the pigs.

Two participants requested environmental sampling including other livestock in the region, plants, other

wild ungulates, and soil microbes for dioxin levels. Two participants requested more thorough epidemiological data, one specifying data on human health outcomes and the other requesting data on human and ungulate health. One person wanted information on pork intake for people in each country. One person requested the full report from the WHO/FAO investigation mentioned in the scenario.

Determining Who is Responsible

The second phase of the survey for Scenario 2 asked participants to consider who or what was responsible for the biological event observed in the scenario. The same pieces of evidence used in the first part were used again for this part of the survey in hopes of seeing if usage of the evidence changed depending on the question. Participants were first asked who or what they thought was responsible and to explain their response. Four participants chose not to answer the question. Of the ones who answered, their response to who or what was responsible can be found in Table 4.10. When the participants were then asked their confidence level, the mean confidence level was 4.73 (standard deviation 2.25), as illustrated in Figure 4.7.

Responsible Party	**Count**
ASF Vaccine Factory	4
Feed	4
Hemera	1
Naturally occurring	3
North Sol Cartel	16
North Sol Cartel and ASF Vaccine Factory	3
Not enough information	6
Nyx	1
Nyx, Hermera, Sol, or a cartel	1

Table 4.10: Responsible party for the biological event in Scenario 2 chosen by participants.

Participants were given ten different pieces of evidence to consider when determining which the responsible party for the biological event in Scenario 2. Each participant had to rank all ten pieces of evidence based on how influential the evidence was in their determination. Only one participant had an identical ranking between each section of Scenario 2. No two participants had identical rankings. The mean rank for each piece of evidence is reported in Table 4.11. A graphical representation of the distribution of ranks for each piece of evidence can be found in Figure 4.8.

A Shapiro- Wilk Test was done to determine if the ranking data fit a normal distribution. The test statistic

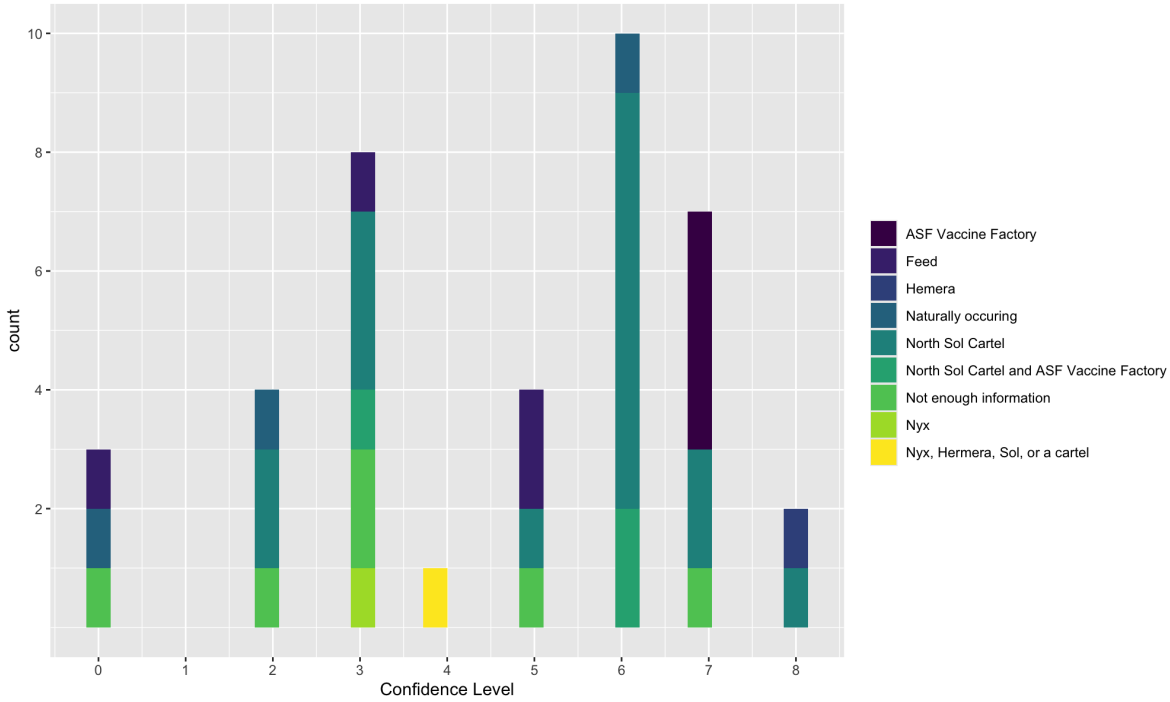

Figure 4.7: Distribution of confidence levels of participants when determining how confident they were in determining the responsible party for the biological event described in Scenario 2.

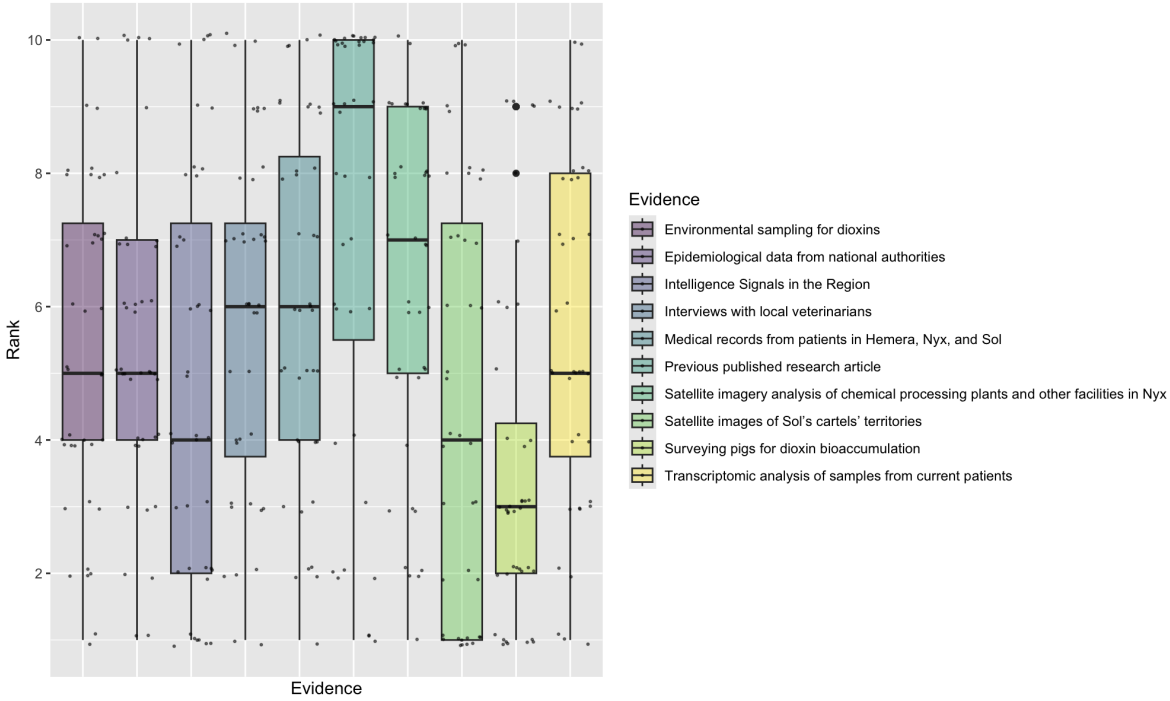

Figure 4.8: Graphical representation of the distribution of rankings of evidence relative influence on determining the responsible party of biological event described in Scenario 2.

Evidence	Mean Rank	Standard Deviation of Mean Rank
Surveying pigs for dioxin bioaccumulation	3.58	2.51
Environmental sampling for dioxins	5.42	2.5
Transcriptomic analysis of samples from current patients	5.68	2.64
Medical records from patients in Hemera, Nyx, and Sol	6.02	2.66
Satellite images of Sol's cartels' territories	4.65	3.14
Intelligence Signals in the Region	4.82	3.06
Epidemiological data from national authorities	5.42	2.33
Interviews with local veterinarians	5.78	2.6
Satellite imagery analysis of chemical processing plants and other facilities in Nyx	6.4	2.6
Previous published research article	7.22	3.25

Table 4.11: Mean rank of different pieces of evidence's influence on determining the responsible party for the biological event occurring in Scenario 2.

was 0.98 and the p value was 1.84 x10^{-5}, so the null hypothesis that the data fit a normal distribution was rejected. Because the data did not follow a normal distribution, a Kruskal-Wallis Test was used to determine if there was a significant difference in mean ranks of the ten pieces of evidence. The test statistic H was calculated to be 43.8 with 9 degrees of freedom and a p value of less than 1.52 x10^{-6}, so the null hypothesis that the rankings for each piece of evidence originated from the same distribution was rejected. At least one type of evidence has a median ranking different from the other pieces of evidence.

A Wilcoxon Sign Ranked Test with continuity correction was used to determine which pieces of evidence had significant differences in their median rankings. The pairings of evidence with significantly different median ranks were determined based on which pairings had a test statistic W less than the p value of 0.005. Test statistics for each pairing are reported in the appendix (Table A4). Seven pairings of evidence were found to have significantly different mean ranks.

Several theories based on the evidence were put forth by participants concerning attribution for Scenario 2. Participants quickly assumed the health outcomes in humans were tied to dioxin accumulation in pigs. One stated, "it is obviously through the food - animal fat accumulation of dioxins." However, participants were split on how dioxins were getting into the human food supply, with the most common theories surrounded feed or ASF vaccines. Pollution was also a pathway suggested albeit as an afterthought in most cases.

Sixteen participants theorized that North Sol Cartel was responsible for the event, with nine of those participants suggesting the mechanism was the vaccine and the other seven theorizing pig feed. Four additional participants thought the ASF vaccine was the mechanism but did not attribute responsibility to North Sol Cartel. Of those who believed the vaccine was the pathway for increased dioxin levels and attributed responsibility to the North Sol Cartel, the intelligence signals and satellite images suggesting North Sol Cartel abnormal behavior and low mortality rates after Year 1 were commonly cited as evidence that supported their theory. When explaining their theory, one participant stated,

> "The simultaneous emergence of ASF in vaccinated pigs in South Sol is consistent with North Sol cartels deliberately tampering with the ASF vaccine as a means of delivering a bioweapon. While the deliberate threat could also have been created by a nation-state, and then been taken advantage of by the cartel, intelligence signals are consistent with a biological event being consistent with a cartel's interests, but not consistent with the interests of any of the other identified players in the region."

Participants who did not recognized ASF vaccine as the mechanism but did not assign blame to North Sol Cartel suggested fault lied with the vaccine manufacturing facility either intentionally or unintentionally altering the vaccine. The timeline, interviews with veterinarians, and prior paper were listed as relevant in their reasoning explanation by these four participants. One participant explicitly blamed Hemera for the event because the vaccine facility was located in that country, though they did not think it was a state sponsored bioweapon.

Participants who thought the feed was the source of increased dioxins reported using the environmental sampling data and the survey of dioxin levels in pigs as key in forming their theories. One participant picked the feed route over the vaccine route due to their analysis of the timeframe given in the scenario,

> "I do not believe this event is related to the vaccine schedule; the vaccine was "given to farmed pigs in the region for 4 years prior to Year 1" with the final dose "at age 1 year". Farmed sows are usually kept for less than 10 parities, and can farrow 3 times a year, so it just doesn't seem plausible to me that you would start seeing effects after 4 years, unless the ASF vaccination campaign has increased in scope in major ways."

Respondents that thought North Sol Cartel was responsible for tampering with feed also noted the intelligence signals and satellite data suggesting abnormal behavior as contributing to their theory. They also noted that North Sol Cartel may be interested in gaining influence and territory or disrupting the legitimate government of Sol through a bioweapon.

Intelligence signals and the prior paper from a laboratory in Hemera helped three participants develop a theory that the North Sol Cartel was collaborating with the Hemera laboratory, which possessed the genetically engineered bacteria that can modulate dioxin synthesis, to alter the ASF vaccine sent to all other pig farms. One of these participants also said the information from the veterinarians contributed to their development of this hypothesis, though at a lesser extent than the intelligence signals. They also said that the low incidence of elevated dioxin in wild pigs and the incidence of disease in North Sol Cartel in Year 1 was indicative of vaccine tampering as the source of the biological event.

One participant theorized that Nyx was responsible via pollution. Their reasoning was that Nyx was slow to agree to participate in the investigation, they had a facility that made dioxin as a byproduct of pesticide manufacturing, and that the event heavily impacted the pork production industry in other countries, but Nyx was the only country in the region that didn't have a robust pork industry but was accused of using bioweapons. The participant thought it could be an accidental environmental exposure from pollution related to the pesticide manufacturing, except for the fact that all other affected countries have an economic hit whereas Nyx doesn't and Nyx wouldn't cooperate in the beginning, and thus concluded Nyx was involved in a state-sponsored bioweapon program. Another participant thought any of the countries could be responsible based on the economic impact of first decreased pork output and strain on the human health system in addition to the impacts of terrorism were people to believe a bioweapon had gotten into their food supply.

There was one participant who mapped out not just what they believed to be the case for determining if a deliberate event occurred and who did it but also included how they wanted to verify their hypothesis,

> "There is a plasmid sequence for a biosynthetic pathway for dioxin. Insofar as there is any case for a deliberate event that event basically has to look like this: North wants Pork-Monopoly. –> Uses Hemera research, and maybe Hemera researchers to make bioagent with dioxin biosynthetic pathway. –> Uses experimental ASF vaccine as initial vector for delivery, but agent is likely somewhat contagious as North has enhanced contagion protocols in place. If

that is the case, pigs with bioaccumulation of dioxin ought to have somewhere in their system, or their feed, sequences associated with dioxin biosynthesis. If not, the whole hypothesis falls apart. So, develop PCR primers for essential invariant components of the dioxin biosynthetic pathway, and verify that the AGENT ACTUALLY EXISTS! Such an agent, regardless of the rest of the scenario, simply CAN NOT be natural... maybe it's an accidental lab-release from the Hemera lab, maybe it's more deliberate but without an isolated agent, epidemiology is insufficient to completely identify motive vs. accident."

Overall Conclusions and Comments on Scenario 2

The primary comment from participants when discussing the appropriateness of evidence presented in Scenario 2 was that the evidence overall was weak. Twenty-three respondents said the evidence was weak, particularly related to the epidemiological data and intelligence reports having too many gaps. Seven participants also mentioned the satellite imagery and prior report to be hard to use in their analyses. Three participants said the transcriptomics data was hard to interpret.

Participants were split when discussing credibility of the evidence. Eighteen said the evidence was less credible because it was vaguer than evidence in Scenario 1 and largely circumstantial and speculative. One participant even said the only non-speculative evidence was the pig dioxin level survey. These respondents frequently said they primarily determined credibility based on how much the evidence relies on speculation verses objectivity. Six participants stated that the evidence was less credible because more came from national authorities, which they did not think could be taken at face value.

One participant gave a breakdown of evidence credibility as follows,

> "The evidence that the biological event is caused by dioxin accumulation in farmed pigs is quite trustworthy: no obvious gaps, clear report of the key features (1, 2, 3, 4, 5, 6). The satellite imaging of cartel territories (7) is difficult to evaluate, since I don't know how reliable it is to draw the given conclusions from satellite imaging. The interviews and associations (8,9,10) are suggestive and consistent with the conclusions, but are too sketch to be definitive."

Ten participants stated that having a multilateral investigation, the WHO investigation, or the UNSGM standards guiding the investigation as positives for credibility of the evidence. These individuals said who

was doing the investigation or how the investigation was conducted (such as how much control the investigatory team had in collecting information) was the key factor in determining credibility.

Participants felt much more uncertain in Scenario 2 than Scenario 1. One stated, “much lower trustworthiness and credibility than Scenario 1, resulting in much higher uncertainty.” Seven participants felt the vagueness of Scenario 2 would be reflected in a real-life scenario; “Just enough data to get participants to bite on a deliberate event but not enough to conclusively point the finger. Often, that’s how it goes in real life. Solving incidents conclusively is easier said than done.” Three people felt the opposite and that real life would be much clearer than the scenario.

Participants were asking for any closing thoughts. Two responses reflected on how these scenarios would compare to a real-life event. One stated, “The data presented is, barely, enough to call for an international team to come and further investigate. International is good because it removes bias. Without an isolated agent, however, it’s just supposition and circumstantial evidence.” Another response was,

> “ I think this, again, shows that peoples’ different perspectives can make it very difficult to know exactly what is an accidental, natural, or deliberate biological event. Unfortunately, there are so many gaps that need to be filled with "reasonable" assumptions, but those assumptions may change based on the information available as well as the composition of the group.”

4.4 Discussion

4.4.1 Limitations

Results from this research should be interpreted with caution due to several limitations, most notably sampling and ability to replicate a true investigation. The sample size of 41 participants is much smaller than the population of people who would be involved in investigating or acting on an investigation of a biological event and is likely not truly representative of the worldwide population. Efforts were made to gain expertise, gender, racial, and cultural diversity in participants of the study. Survey materials were only made in English and therefor all participants had to speak English, biasing results. Additionally, response rate and agreement to participate was higher among experts from the US and UK, likely in part due to name recognition of the US based study team. The scenario and survey were developed by the US research team and

therefore influenced by cultural biases from the US perspective. Future research in this area would benefit from collaborations with researchers from other countries and cultures to reassess the scenario and improve participation among non-American participants.

This study was exploratory. Results must be validated with reproducibility in other studies. Different scenarios and participants could generate different trends in responses. Additionally, the methodology used in this study attempts to replicate a true investigation of a biological event. However, due to time constraints on participants the quantity of data that would be expected for an investigation, it is not practical to attempt to perfectly replicate an investigation. The methodology employed for this study was designed to extract key components of an investigation report, such as identifying who collected or analyzed the data, and representative pieces of evidence to compare biological samples versus more traditional intelligence and evidence. Limited detail about the evidence was provided to avoid over-explaining or interpreting evidence to limit bias imposed from how the study team wrote the scenarios. Future work that creates a more realistic and immersive experience could provide more robust and representative results.

This exercise was designed to provide information that could be broadly applicable to a real-life situation. However, participants may react differently when a situation is real and not an exercise and the impact of world events and politics will inevitably impact how individuals view any given situation. Ideally, the way in which people would approach attribution would not change drastically in a real situation or in scenarios with different details. Conclusions may vary but how an expert goes about examining evidence and their expectations for how an investigation is conducted should be stable. A repeat of this experiment could provide valuable information concerning changes in attitudes on these topics.

4.4.2 Pilot Study

The results from the pilot study revealed several key findings that were important for guiding further research efforts. Results from the pilot studies of both scenarios suggested that the chosen methodology was appropriate for testing how stakeholders would be assessing and interpreting evidence during a biological event. The scenarios and survey were produced data that could feasibly be analyzed and were well-received by participants. Given that pilot participants reported uncertainty in their conclusions and there was diversity in answers about who was responsible for the events, the scenarios were not altered as the goal was to avoid

the scenarios being obvious. In real life events, there will be great uncertainty in what has happened, and people will have to assess evidence, credibility, conclusions, and reporting on the incident for themselves. After the pilot was conducted, the authors were confident they had succeeded in replicating a situation with an appropriate amount of uncertainty.

4.4.3 Scenario 1

Scenario 1 focused on the death James Salle. Participants were asked to respond to the scenario in two parts, the first part focused on exploring if Salle was deliberately killed with a biological agent and the second part focusing on who may have been responsible for killing Salle. Participants were more uniform in responses for part one compared to part two.

Assessing Laboratory Evidence

Biological evidence is a type of physical evidence, which is routinely assumed to be more objective and therefore reliable than testimonial evidence (172). Evidence types that required a laboratory for analysis were typically ranked higher than evidence from more traditional investigatory processes that did not involve the laboratory, with four of the five laboratory-based evidence types being ranked higher than any of the more traditional forms of evidence based on the aggregate rankings. As was noted in the subsequent question, some participants felt biological evidence "couldn't lie" and was therefore more trustworthy than other types of evidence, like interviews and witness statements. This may account for the why the biological evidence was generally ranked higher than non-biological evidence. However, other types of evidence that are also generally considered to be "the truth" (in a pre-generative AI world), such as surveillance footage, had a lower aggregate ranking than other types of evidence that can be considered unreliable, like witness statements. The participants were asked to rank the evidence based on how influential it was in relation to the other evidence. "Influential" was picked because it could include considerations such as trustworthiness, reliability, accuracy, and importance for decision making. Open ended questions following the initial ranking asked participants to discuss trustworthiness and tease apart what was considered by the participant in determining influence on decision making.

Evidence included in the scenario was given a brief description. For evidence requiring laboratory anal-

ysis, a quick description of the results was provided. Sequence similarity, number of copies of a particular gene, mutations in a gene, and descriptions of clade groupings were included to assess how participants interacted with these types of genetic analyses. Rabies virus is a negative sense RNA virus that relies on an RNA dependent RNA polymerase (RdRp) to replicate. The rabies virus RdRp has a high mutation rate, leading to the virus existing as a quasispecies (173). The percentages of similarity provided to participants were designed to be plausible for a quasispecies producing virus, as were the descriptions of clades and the number of gene copies. Participants had different levels of familiarity with genetic analysis. The scenarios did not include a primer on how to interpret results. Instead, the type of genetic analysis included was designed to be simple enough for people with high school biology to be at least somewhat comfortable understanding. In a real-life scenario, levels of education and backgrounds will differ. Participants were welcome to use the internet or textbooks as a resource to better conceptualize the evidence provided, just as decisions makers will be free to seek their preferred advisors.

Sequence similarity is reported by a percentage and was considered as the most straight forward genetic analysis provided because percentages are used in many fields. Nucleotide sequence similarity percentages were provided comparing sequences from the dart to samples taken from Salle (99.2%) and between analyses of samples from Salle (98.7%). The degree of similarity between sequences was chosen to not be a perfect match to reflect differences that could be due to degradation of the sample over time, differences in sequencing protocols and reagents, and how the viral population itself could evolve within different host tissues. The relatively small differences were meant to see if these were big enough differences for participants to question if the dart was the source of rabies for Salle or if the investigatory team compared to initial medical examiner samples were suspect or one sample potentially less valid than the other. Interestingly, the metagenetic analysis from the dart was interpreted very differently amongst participants. For some people, it was the most important and clear evidence that a deliberate event had occurred. For others, it was at best ambiguous and at worst meaningless because it was impossible to determine directionality of transmission: did Salle get rabies from the dart or did an already infected Salle get rabies on the dart? Additionally, the similarity in sequence between the dart sample and postmortem samples was highly impactful for some people with some participants reporting not only was the similarity enough to determine a deliberate event had occurred but also to attribute the crime to one suspect. Others thought it was odd the sequence similarity was

"so low" between the dart analysis and postmortem samples, while even more participants didn't know how to interpret any of the sequence similarity data. Given that these were the two most highly ranked pieces of evidence in part 1, it is surprising that there is such diversity in how people used them in their analysis. This highlights how one piece of evidence that some think is straightforward and critical might not be to others. A potential vulnerability of an investigation could be such pieces of evidence where there is great emphasis on a given piece of evidence but vast differences in how people interpret the evidence. Such differences in interpretation could be major candidates for exploitation for misinformation or disinformation. In part 2, some participants said the sequence similarity alone was enough to determine Smith was responsible. Other people stated they didn't have a good idea of how to interpret the percent similarity numbers between sequences, and they typically ranked such evidence lower in this section, unlike in part 1.

However, in the next section of the scenario, when discussing samples taken laboratories connected to suspects, nucleotide similarity of 86.1% to 96.9% and 83.2% are provided. These lowers percentage similarities did make some participants, including some who did not identify as scientists, question if there was a true match between the rabies on the dart and the samples taken from laboratories. 86.1% to 96.9% was meant to indicate there were many samples of rabies virus in the laboratory that had either come from different sources or been passaged extensively enough cause such a difference in sequence similarity. In either case, this would suggest a laboratory that has many vials of rabies available and that some vials contain rabies virus very similar to what was found to infect Salle. The other facility was reported to have a rabies lab, but the highest degree of similarity was only 83.2%. Some participants did report this 86.1% to 96.9% range of similarity to be influential in their decision making, especially compared to the 83.2%.

Scenario 1 included evidence that found there were two copies of the glycoprotein gene found in rabies virus population sampled from Salle after death. In the evidence description, participants are told the scientific literature suggests multiple copies of the gene could increase infectivity and that two copies of the gene have not be found in a rabies isolate. The scenario also stated there were mutations in the duplicated gene at 4 sites and that these mutations were reported to enhance viral pathogenesis in the literature. In their analysis, most participants assessed the mutations to mean there had been genetic engineering of the virus that killed Salle (than thus intention to harm him). Interestingly, the presence of mutations in both genes, but not the presence of a double gene itself, was considered to indicate genetic engineering by most

participants. Mutations may be better understood by participants compared to gene duplications and the capacity for viruses to handle multiple copies of a gene.

One piece of evidence included a description of a phylogenetic tree comparing samples taken from Salle to samples taken from a laboratory connected to a suspect. The isolates from the laboratory grouped together while Salle samples created their own clade. Only one participant mentioned the phylogenetic tree when describing how they analyzed their evidence, suggesting the concept of a phylogenetic tree is less impactful for a broad audience compared to the percent sequence similarity. The evidence including the phylogenetic tree was paired with the 83.2% sequence similarity, so the phylogenetic tree alone can't be compared to these other types of genetic analysis to assess how the different types of information ranked in consideration for participants. **However, the participant above noted that while most people do not have the expertise needed to fully read a phylogenetic tree, they can be very useful for attribution, especially when considering a phylodynamic tree.**

Importantly, a single number for sequence similarity will not be an adequate description for comparing two viral populations. It is one of the simplest metrics, and an approachable one for non-scientists, but it does not adequately describe the viral quasispecies or indicate the impact of the differences in the viral population. In a real investigation, growth curves, phylogenetic trees, and detailed descriptions of the quasispecies traits will be needed. However, this information will not be nearly as ingestible to the public as a sequence similarity. If there is a case where sequence similarity could point to one conclusion, but this other data more comprehensively suggests an alternative conclusion, this could be an opportunity for exploitation by conspiracy theorists or supporters of the first conclusion.

Environmental samples were included in the evidence list because environmental sampling has been used to determine points of origins for outbreaks in the past. The scenario says entrances, exits, handrails, and the stage were swabbed. These swabs were used for PCR testing to detect rabies virus fragments. However, the scenario specifies that testing took place 43 days after Salle died. The delay in testing was included because rabies was not known to have infected Salle immediately after the dart hit him. This evidence was included to show there was not a continuous presence of rabies virus at the venue. Unsurprisingly, this environmental testing was ranked the lowest of the laboratory derived evidence. Participants stated they didn't think the environmental sampling from the facility was done soon enough after the event. Had the environ-

mental sampling been done sooner after the attack, some participants may have found it more impactful for their assessment. However, it was the only type of laboratory derived evidence that multiple participants suggested was "useless". In open ended questions, multiple respondents suggested environmental sampling would not have any place in such an investigation. This attitude may stem from an ongoing controversy at the time of this experiment related to environmental sampling of viruses (174).

There are different types and reasons for environmental sampling. Such sampling can be done on water, air, soil, or other surfaces in a built or natural environment. Sometimes sampling is done for disease surveillance, such as wastewater surveillance for a particular pathogen causing disease in that community. While environmental sampling for public health surveillance has some drawbacks, such as concerns for privacy, it is not particularly controversial among politicians or the public in Western countries. Alternatively, since the origin of COVID-19 became a highly politized topic, environmental sampling for research purposes or "virus hunting" has become controversial, with some people suggesting such work doesn't provide enough benefit for the cost and potential risks involved. While the example of environmental sampling given in this scenario was not for research, some participants did suggest any environmental sampling would be useless and a safety risk. The conflation of all environmental sampling and associated controversy could be a problem for investigations in the future. This is another area where an investigatory team will need to be cognizant of the general opinions and attitudes at the time of the investigation to avoid their work being inappropriately targeted or rejected.

Participants overwhelmingly found biological evidence the most credible, followed by film/photo evidence, then statements from individuals. However, much of the biological evidence was genetic analysis that many participants stated they did not fully understand yet still ranked it higher than other evidence they did understand. This could be a problem in a real investigation if the genetic evidence is also the most divergent in terms of how people interpreted it. Many people inherently trusted genetic information more than other types of evidence, taking it at face value without fully understanding the results or the methods required to get the results. Genetic information was also the most divergent as far as credibility and interpretation from other biological and non-biological evidence. This contradiction of high perceived credibility but low understanding makes it another vulnerability for exploitation by those wishing to undermine the investigation.

There are different types and reasons for environmental sampling. Such sampling can be done on water, air, soil, or other surfaces in a built or natural environment. Sometimes sampling is done for disease surveillance, such as wastewater surveillance for a particular pathogen causing disease in that community (175). While environmental sampling for public health surveillance has some drawbacks, such as concerns for privacy, it is not particularly controversial among politicians or the public in Western countries. Alternatively, since the origin of COVID-19 became a highly politized topic, environmental sampling for research purposes or "virus hunting" has become controversial, with some people suggesting such work doesn't provide enough benefit for the cost and potential risks involved (176). While the example of environmental sampling given in this scenario was not for research, some participants did suggest any environmental sampling would be useless and a safety risk. The conflation of all environmental sampling and associated controversy could be a problem for investigations in the future. This is another area where an investigatory team will need to be cognizant of the general opinions and attitudes at the time of the investigation to avoid their work being inappropriately targeted or rejected.

Participants overwhelmingly found biological evidence the most credible, followed by film/photo evidence, then statements from individuals. However, much of the biological evidence was genetic analysis that many participants stated they did not fully understand yet still ranked it higher than other evidence they did understand. This could be a problem in a real investigation if the genetic evidence is also the most divergent in terms of how people interpreted it. Many people inherently trusted genetic information more than other types of evidence, taking it at face value without fully understanding the results or the methods required to get the results. Genetic information was also the most divergent as far as credibility and interpretation from other biological and non-biological evidence. This contradiction of high perceived credibility but low understanding makes it another vulnerability for exploitation by those wishing to undermine the investigation.

Willingness to rely on the biological evidence, specifically the sequence similarity, when determining the type of event was higher compared to when assessing a guilty party. It was not clear from participant responses why some changed their willingness to use sequence similarity data between parts 1 and 2, or even if they acknowledged they had done so. Such a switch could be caused by the different nature of these questions and implications of deciding. The first question was asking what happened but didn't place blame

whereas asking participants to decide who was responsible would likely lead to severe consequences for that individual. Participants may have been less willing to use evidence they didn't understand when there was such high potential consequences associated with a decision.

Determining if there was an Assassination via Bioweapon

The first part of this scenario asked participants to determine the nature of Salle's death. First, they had to consider if he was deliberately killed. Second, they had to decide if a biological agent was used to kill him. Of the participants that thought a deliberate event had occurred from the beginning of the exercise, there were two main approaches in reaching that conclusion. Some participants were looking for lack of a natural transmission pathway to indicate a deliberate event while others were looking for evidence of nefarious intent. Those that were looking for nefarious intent often used terms like "smoking-gun" when referring to the genetic data and heavily relied on that data compared to all others. Participants who seemed to be looking to rule out natural causes first before considering a deliberate cause of death were more likely to mention more types of evidence in their written responses and more likely to rank all the biological evidence above other types of evidence. This bifurcation in how participants were approaching deciding could impact not only what evidence is collected but also confidence in a conclusion. Participants who determined there was a deliberate event were more confident in their conclusion than those selecting the inconclusive option. One participant looking for a lack of a natural transmission pathway and who ultimately selected "inconclusive" said they were trying to "prove a negative." Proving a negative is much harder than proving a positive claim because it requires more exhaustive searching. If some experts are going into attribution expecting an investigation to prove a negative, they may never accept the results of an investigation because it will never be thorough enough. Understanding and managing expectations from both investigators, decision makers, and the public will become a critical responsibility.

When determining if a biological event is naturally occurring or caused by a human behavior, genetic analysis may not be the most helpful evidence. Genetic information will change as organisms replicate and importantly, evolve. While scientists know of factors influencing evolution, such as environmental pressures, and can guess how those pressures will impact the population, we can't predict the exact genetic changes. Scientists can use site directed mutagenesis to insert specific mutations to get a known outcome. Scientists

can passage a microbe under specific environmental conditions to get a population with a specific phenotype, such as resistance to the mutagen. Scientists can also introduce a lot of random mutations. However, there is no way to know the exact genetic changes that will occur in response to a given environment without prior study. If there is an artificial nucleic acid analogue present, it there will be strong evidence for genetic engineering.(177) A sequence that appears to have a full gene from a different organism or has markers of using a gene editing technology could also be indicative of genetic engineering. But there will be many situations where anomalies found in a genome could be naturally occurring and you can't determine if it has been genetically modified or not.

Looking at phenotypic expression may help when genetic analysis isn't definitive. Understanding the baseline growth kinetics, infectivity, pathogenesis, and other characteristics of a pathogen and comparing to samples in question can shed light on if there has been modification to the agent. However, changes could still be a product of evolution. Additionally, while it may be possible to collect genetic information for sequencing, there may not be samples for passaging to study the phenotype. Passaging itself will cause evolution of the pathogen, so even phenotypic expression will need carful and restrained analysis.

Even if there is evidence of genetic modification, that may not correlate to nefarious intent. Bacteria and viruses are genetically modified in labs to create vaccines and treatments or to study the diseases they cause. Sometimes bacteria are modified to produce a product like spider silk or something people can eat. These modifications are not nefarious and there are typically extensive biosafety measures in place to prevent these organisms from harming someone. If someone was infected by one of these organisms, samples would indicate genetic modification but not intent. Genetic analysis alone cannot determine intent. Other types of evidence would be required to make such a determination.

Interestingly, one person said, "while the rabies virus was altered, I would not necessarily call it a bioweapon. The perpetrator used a strain that was more infectious to lessen the chance of survival, and it was a deliberate attack. That does not necessarily mean it was a bioweapon. It was a targeted attack on one individual using a biological agent." Such nuance in terminology may impact how decision makers would react to a report. Using "bioweapon" in a situation like this scenario could be seen as sensationalist or fearmongering, leading to less willingness to act on the information. Using the term "bioweapon" could also change how international law and norms are applied to the situation.

Determining Who was Responsible for the Assassination

In the second part of Scenario 1, participants were asked to find the party responsible for Salle's death. The evidence for this section was written to cast an equal suspicion on the two main suspects while also leaving room for the possibility of it being neither suspect. Roy Smith was written to have a clearer motive and being more aggressive with police but does not currently work in a laboratory working with rabies. Jim Little was cooperative with the police, works in a facility with a rabies laboratory, and has some suspicious behavior going into that lab but lacks a motive. Smith had the most votes with higher confidence than Little or the "another suspect option".

From the ranking data, scientists more frequently choose the samples from labs in their top three pieces of evidence compared to others. Familiarity with viruses and sequencing may be leading scientists to value these types of evidence more whereas people less familiar with the technology could be more reliant on other types of evidence that are more familiar to them. In contrast to the scientist group, those that identified as being in one of the security-related areas of expertise chose Jim Little at a much lower rate than those not in a security related area of expertise. In their descriptions of why they picked a particular subject, participants identifying as an expert in a security related field exclusively mentioned motivation as part of their decision making. This difference in response could be due to training in investigations by fields of practice. It would be interesting to see how conclusions would change if given the opportunity to collaborate on the investigation.

Scientists were more likely to choose a named suspect than other areas of expertise, particularly choosing Jim Little. Scientists picked Jim Little at a higher rate than non-scientists (36.8% versus 9%). Conversely, scientists picked Roy Smith at lower rates than non-scientists (26.3% versus 45.5%). Little was described as a scientist. Many participants who identified as scientists described Little as having expertise to engineer a virus, access to a stock virus, and access to the tools and knowledge needed to take the samples from his workplace that were 83.2% similar to the Salle post-mortem samples and create the virus that killed Salle. Scientists picking the scientist suspect at higher rates than other professions could be due to scientists being exposed to biosafety training at a higher rate than other professions. Even the most basic biosafety trainings will touch on potential misuse of biological agents or chemicals from the lab. Alternatively, scientists are more familiar with how laboratories run and may have been more suspicious of the unauthorized entry

into the rabies lab by Little compared to non-scientists. Scientists may have been quick to discount Smith as responsible for the assassination because they didn't think he had the training required to do the engineering, as multiple participants who identified themselves as scientists stated in their reasoning. One participant who is a scientist mentioned they thought Little was likely to have access to and the skills to find and interpret literature that could point to which mutations to engineer whereas they didn't think Smith had the same access and skills. Further work is needed to understand if there is a bias by scientists towards other scientists and if so, how this bias could impact an investigation. As DIYbio becomes more popular, biases against people who don't have formal scientific training should also be explored. In this scenario, some scientists were quick to dismiss Smith for not being a professional scientist.

Confidence in conclusions was much lower when considering who was responsible for throwing the dart compared to determining if Salle had been intentionally killed. A similar phenomenon in seen when comparing variances amongst rankings between the two sections of Scenario 1; participants were more confident and more uniform in thinking when determining what had happened compared to determining attribution. Based on participant statements, the lack of confidence in a conclusion for this section is caused by confusion or dissatisfaction with the evidence.

Participants were split on the impact of facility inspections. Several participants ranked the university laboratory inspection as one of their top two pieces of evidence while others ranked it in the bottom two, stating the inspection wasn't adequate for drawing conclusions. Alternatively, participants were widely in agreement that the social media analysis was a weak piece of evidence and the interviews with suspects were typically ranked low, though there were some participants ranking one of the interviews in their top two positions. The difference in attitudes towards facility inspections could be attributed to level of familiarity with laboratory spaces; those with more familiarity with laboratories wanted more information from the inspections and more samples collected. Social media is more widely known by the public which could make it easier for all participants to judge for quality and impact.

Participants stated trust in the investigatory team and process as important factors for this scenario. Interestingly, several participants said this was more impactful for determining attribution than determining if the event was deliberate. This could again be due to the potential consequences to an individual if determined to be at fault.

4.4.4 Scenario 2

Scenario 2 featured a public health crisis of ambiguous origin in the fictional countries of Sol, Nyx, and Hemera. Participants were asked to determine first what the nature of the event was and then who may be responsible for the event. Scenario 2 was designed to mimic a nebulous event; no one knows exactly when it started or why, but overtime it becomes clear that something is happening. The evidence for this scenario was deliberately created to be more vague and high level to replicate a long-term event that is not initially detected. Scenario 2, unlike Scenario 1, is a collection of disease cases across a region and time. There isn't one specific case to focus on in this scenario and participants were forced to analyze intentionally vague evidence to assess the overall situation. Confidence in conclusions drawn from the provided evidence, and satisfaction with the evidence itself, was lower with this scenario compared to the previous.

Determining the Nature of the Event

Most participants (35) selected that the event described in the scenario was accidental or deliberate while six participants chose naturally occurring. However, confidence in this conclusion averaged low at 4.8. Those choosing a natural event over man-made were less confident than their counterparts with an average confidence of 3.3 compared to 4.6 for accidental and 5.4 for deliberate. The challenge of "proving a negative" was mentioned again for this scenario, however it was mentioned in relation to both natural and accidental. For both cases, some participants described their thinking as if they defaulted to natural first and then if enough to suggest there was something not totally natural, then an accidental cause before considering deliberate. In this scenario, some participants considered the fact that there was severe disease in humans that could not be traced back to a known source was sufficient for concluding the event was not naturally occurring. While other participants thought the human disease could be explained by natural factors, even if such factors were not readily identifiable. Multiple reasonings suggested a participant would not move from natural or accidental unless there was evidence to not only prove deliberate but also disprove natural or accidental. There may be an inclination to stick to natural or accidental because such events happen more frequently than deliberate events. Alternatively, there may be a hesitation for choosing deliberate due to the potential consequences to an individual or a fear of being labelled a conspiracy theorist. Some participants stated that independent of evidence, they would default to naturally occurring automatically and only

once natural was disproved would they consider accidental and deliberate. This again makes a high bar for attribution in a real-world scenario and indicates managing expectations will be critical.

There was little significant difference in mean rankings. The survey of dioxins in pigs was statistically different from the bottom ranked evidence, satellite imagery and previous research, but there was otherwise no significant difference in mean rankings. Participants repeatedly stated how inadequate they found the evidence and requested many types of evidence that they recognized would be impossible to provide.

Determining Who is Responsible

Several different potential responsible parties were identified by participants, all with relatively low confidence. In Scenario 1, participants had to choose between three options, two of which were named individuals. In Scenario 2, participants were free to pick any actor in the scenario as being at fault. North Sol Cartel had the most votes with sixteen individual votes and four more votes in combination with the ASF vaccine factory. Whereas Scenario 1 had specific evidence that evenly pointing to different people, Scenario 2 is highly ambiguous. Giving participants less direction on possible suspects likely contributed to the breadth of suspects identified and contributed to the overall lower confidence in the conclusion as compared to Scenario 1.

Mean rankings differed between sections of this scenario, most notably with the satellite imagery having a wider distribution in mean rank. Participants didn't think this imagery helped determine the nature of the public health crisis but did think it helped determine who could be responsible. Satellite imagery could show changes in behavior or suspicious behavior, which could be used to determine motive. Interestingly, despite being helpful in determining fault, it didn't have a major relative impact on determining if a deliberate event occurred. This suggests people were not considering who could have done it when considering if something was intentionally done.

The most common comment from participants in both sections of Scenario 2 was a request for more information. Participants did not like the uncertainty, and some reported getting frustrated and uncomfortable with Scenario 2. There may be an unconscious expectation to know to nature of a major event and when that doesn't happen, it is uncomfortable. Several psychological studies have addressed the need humans have for control and understanding of their environment (178–180). Studies also suggest that during the heightened stress of a pandemic or outbreak, desperation for this control and understanding will increase (181). Such

desperation will make people more amenable to mis- and disinformation which will make it harder for an investigation to move forward. It could also affect how willing people are to trust an investigation report, use the report for decision making, and the safety of those involved in the investigation.

Participants sometimes would not engage with the evidence provided and said they couldn't do anything without more or better evidence. While it would have been impossible to have a full-scale investigative report for an exercise of this size based on a fictional scenario, there was intentionally only a few pieces of high-level evidence included. For many events, especially those with international geopolitical considerations or taking place over long periods of time or conflict-ridden regions, it is going to be hard to collect evidence. In a real investigation, there would be considerably more information provided to decision makers and the investigatory team, but there would not be all the evidence someone wanted. Investigators would have to do the best they could with whatever information they could collect. Decision makers and the public would have to accept the limitations to what could be collected.

4.4.5 Implications for Research

Overall, participants felt more confident making conclusions about Scenario 1, which was anticipated since it was the more specific event. However, something ambiguous, like Scenario 2, is the bigger threat. Scenario 2 reflects the start of many naturally occurring events; something is happening, but it is undetected for some length of time and there are gaps in data. Naturally occurring threats are more common that anthropogenic caused events, but it is easy to find ways to blame a group when there are gaps in data and tensions are high. Scenario 2 also affected more people and multiple species. With more people affected, there is a higher quantity of data, but also more imperfect data. Sometimes imperfect data is conflated with intentional misleading, which can be exploited by bad actors. Situations like Scenario 2 would rely on more sources for data, each of which could be questioned for credibility or targeted. Scenario 2 requires more inter-professional work to mitigate the threat and investigate the event. Ambiguous events like scenario 2 are bigger, scarier at the population level, harder investigate, and have a greater impact on geopolitics and everyday lives. Scenario 1 was much more straight forward but there were still large differences in how participants responded to attribution in that scenario. Situations like Scenario 1 will be important for learning and building protocols and best practices for ambiguous situations like Scenario 2.

All sections had large differences in how people interpreted the same pieces of evidence, even within the section with the most agreement among participants (Part 1 of Scenario 1). This was most evident when looking at the genetic sequences. Some people exclusively used the genetic analysis data to make their determination while a small set of others said they didn't use it at all. Both camps had participants that said they didn't understand the genetic evidence. Several people said they heavily relied on the genetic analysis because it was entirely objective and the most accurate evidence. However, only one participant questioned what methods were used or quality control of the sequencing and only two others mentioned that the genetic analysis could be wrong. Participants repeatedly placed high importance on biological samples, especially genetic analysis, regardless of if they understand it or not, suggesting this type of evidence will be expected and critical in an attribution context. However, if genetic evidence is routinely found to be the most trustworthy and impactful evidence but there are widespread differences in interpretation of that evidence, there could be problems when it comes time to draw conclusions as a group. This phenomenon was seen during the COVID-19 origins debate when there were people pointing to parts of the SARS-CoV-2 genome as evidence of genetic engineering or the sequence similarity between different samples. There was an agreement that genetic information was important, but how people interpreted the information differed. People with and without training on viral genetics were confidently claiming opposing conclusions based on the genetic information. Just as demonstrated here, people will use evidence without understanding the science behind the evidence. People don't know what they don't know. This experiment didn't allow people to work in teams, but future work exploring how people of different backgrounds work together to assess the situation would be enlightening. If Scenario 1 was run again but participants were paired security expert with scientist, the groups with the most divergent conclusions regarding attribution, results might change. Results could also be impacted based on if the pair have a prior, respectful relationship or if they are strangers.

Understanding and managing expectations of an investigation and evidence from investigators, decision makers, and the public will become a critical responsibility if a biological event occurs. Clear communication about what is possible, what gaps exist and why, and who is doing the investigation will be vital to protect the integrity of the investigation and gain credibility. There isn't going to be perfect data available nor all the data some would desire. There will also be the problem of "proving a negative" and people will require different levels of prove to disprove their pre-conceived defaults.

Evidence analyzed using laboratory methods requires someone with specialized training in science do the analysis physically in the laboratory and analyze results to draw a conclusion. The scientist is an integral part of this process as that evidence is reliant on those skills sets to be useful. While participants in this experiment were overall mostly willing to trust that the scientist analyzing the data was correct and a reliable source, the same may not occur in a real-life situation. Given widespread mistrust in science today, people may not be willing to trust not only scientists commenting on data that is publicly available, but trust in this nature of evidence may erode.

Chapter 5

Developing a Conceptual Framework for Biological Attribution

5.1 Introduction

Investigations into potential weapons of mass destruction use will be highly politicized and complex operations. The officials responsible for initiating an investigation and those tasked with conducting the investigation will be under high global scrutiny. This will be especially apparent for biological agents, especially if the suspected agent is still circulating in human or animal populations. Therefore, it will be critical for officials to prepare for such investigations before they begin. Preparation should include creating guidance documents and protocols, as well as including as many contingency plans as possible. During an investigation, the investigation team will definitely need to adapt as the investigation proceeds. The guidance documents should help guide the team through a challenging and likely changing landscape. A conceptual framework illustrating the factors that will influence an investigation can help in the development of such guide documents.

Biological events can encompass a wide range of events. The type of agent used (bacterial, viral, toxin), the delivery mechanism, the intention of the perpetrator(s), and even the nature of the event (accidental, deliberate, or naturally occurring) can vary. The guidelines must be specific to the needs of biological events, while being flexible enough to cover the range of possible scenarios. The fundamental stages of the

investigation, such as detection/initiation, gathering evidence, analysis, and reporting findings, should not change over events. Understanding the factors influencing each stage and identifying how these factors can evolve over the course of the investigation or differ between investigations can help shape the road map for investigations.

Results from the prior two aims and a literature review are used to identify, categorize, and map factors that would influence an investigation of a biological event. A conceptual framework for theoretical biological investigations is proposed that incorporates a comprehensive set of factors and the relationships between these factors. The goal of this framework is to aid **responders and agencies with an investigatory mandate** in planning for a biological investigation.

5.2 Methods

In Chapter 3. a list of factors that influence biological attribution and investigation of biological events was identified. Chapter 4 began to assess how these factors could interact with each other in an investigation. To supplement these findings, a literature review was conducted to assess other factors that could impact a biological investigation. Other types of WMD investigations were included in the search, as were studies assessing the use of biological evidence for non-WMD investigations, such as analyzing bodily fluid for metabolites of illegal drugs. A qualitative scoping literature review is used to obtain a full list of factors that could impact a biological investigation.

A preliminary literature search was conducted to identify existing frameworks for investigations. Frameworks for epidemiological investigations and chemical and nuclear event investigations were found, but no frameworks specific to investigating deliberate or accidental biological events were identified.

The scoping literature search was conducted in October 2024. The existing literature was searched using PsychInfo, Web of Science, and PubMed. Keywords for the search were obtained based on the preliminary literature review and include: [(chemical weapon) AND (investigation)], [(chemical weapons) AND (attribution)], [(nuclear weapon) AND (investigation)], [(nuclear weapons) AND (attribution)], [(biological weapon) AND (investigation)], [(biological weapons) AND (attribution)], [(biological forensics) AND (investigation)], and (microbial forensics). This approach identified 3,151 unique articles published between 1992 and 2024. Article titles were screened to ensure relevance to the investigation of crimes. Articles

exploring genealogy, identification of remains, or perpetrator psychology were excluded. This left 2,284 articles for abstract screening. Articles identifying specific biomarkers, vaccine candidates, or metabolites were excluded. The articles included detailed challenges with sample storage, quality, or transport; describing investigations of radiological, biological, chemical, or nuclear events or materials; or describing methodologies for analyzing biological, chemical, radiological, or nuclear samples collected from the environment or a crime scene in the context of an investigation. A total of 920 articles were included from this pool. The content analysis of the articles included assessing the articles for significant relevance to the framework, leaving 134 articles. Articles were categorized for thematic coding which was combined with the coding done in Chapter 3 of this dissertation. The model was built using the codes from these combined sources.

5.3 Model

Each phase of an investigation will be affected by technical, policy, and social considerations. Based on the results of the prior aims, we have a list of characteristics that will impact the investigatory process. Figure X lists these considerations.

Factors identified in Chapter 3 were supplemented with the results of the literature review to build a framework for biological attribution and investigations. There are three main components to the framework: contextual factors, attribute capabilities, and the investigation strategy. Each of these main parts of the framework interacts to enable attribution and biological event investigations. The framework is illustrated in Figure 5.1.

5.3.1 Contextual Factors

Five main contextual factors were identified for biological attribution: governance/policy, ethics, legal, logistics, and politics. Each of these aspects will determine how an investigation is shaped and should be considered when planning or training for an eventual investigation. Ultimately, any of these factors could lead to the failure of an investigation. Importantly, each of these contextual factors can be considered and at least partially addressed with adequate planning before an investigation is initiated.

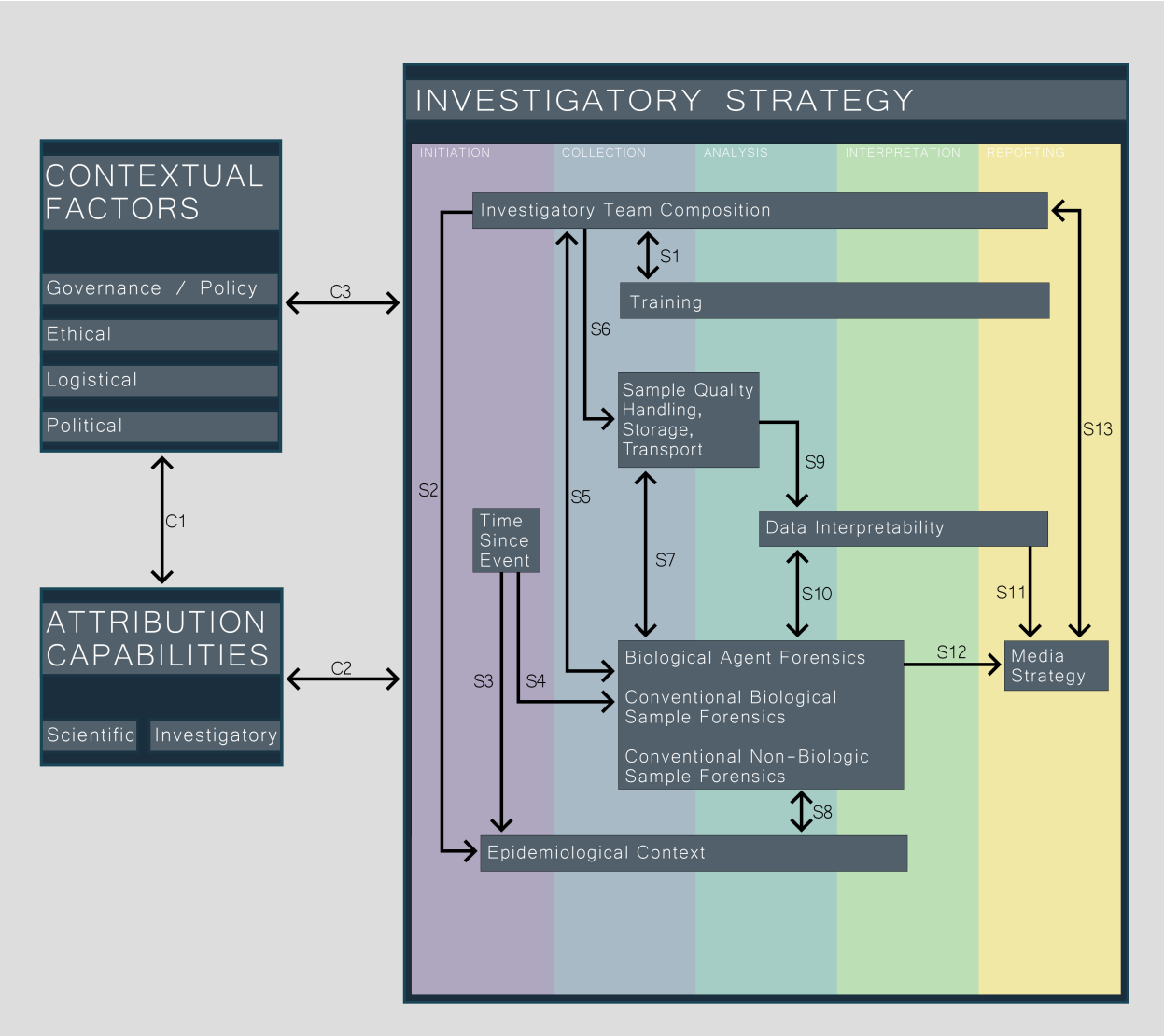

Figure 5.1: Proposed framework for biological event investigations and attribution.

Policy

The governance framework within which an investigation will be conducted will determine who has authority to conduct the investigation, the criminal system in which the case will be tried (if applicable), the standards for evidence gathering, and many other aspects of the investigation. Protocols, procedures, investigators, and available resources will also depend on the policies of the relevant jurisdiction. For example, if a biological event happens within the US and cases occur across state lines and there is evidence to suggest it is a deliberate event, then the FBI will have the leading role in the investigation, and the WMD coordinators for the affected jurisdictions will likely lead the investigation for the FBI. However, if the event crosses state lines but does not have signs of a deliberate event, the CDC will coordinate with state health departments to lead the investigation and respond to the event. If the event appears deliberate, but is contained in one state, then the state's investigation bureau could lead the investigation with the support of the FBI in certain circumstances. In each instance, there would be slightly different protocols and procedures in place. Larger differences in procedure would be highlighted in the case of an international event, when there could be multiple countries affected, each with their own unique procedures, resources, and policies regarding evidence gathering, sharing, and analysis. Preservation of evidence under international legal standards, cooperative agreements, multilateral organization input, international norms and regulations, intelligence sharing allowances, and judicial readiness will all affect multilateral investigations. International norms will also play a key role in multilateral investigations.

The impact of different policies could have far reaching consequences. For example, differences in protocols for evidence gathering could greatly impact downstream conclusions. If one jurisdiction's protocols dictate gathering samples in triplicate by using three different swabs on the same surface while another jurisdiction's policy is to use one swab that is then cut into thirds, there will be differences in results, and it could be difficult to assess how impactful such differences will be on overall findings. Differences in equipment could similarly have impacts that are difficult to quantify.

Multi-jurisdictional planning before an event will be vital for assessing differences in policies and either streamlining policies or addressing where such differences could cause problems for an investigation. If an investigation is started, it will be critical to appoint someone to make sure all policies are being followed appropriately and to list where differences have affected the investigation.

Ethics

Ethical considerations will be key when planning and conducting an investigation. Questions may arise about consent, especially if samples are needed from a local population to check for exposure when they are not exhibiting symptoms currently. Biases could impact the interpretation of results. Lack of understanding for a particular culture could lead to incorrect assumptions. There could be privacy issues. Environmental sampling, including wastewater sampling, which may be used in an investigation, comes with privacy concerns for local populations as personally identifiable data could be generated. If such data are generated, there will need to be built-in protections for handling, storing, and reporting such information. If the expectation is that an investigatory team will hand over all the data generated in the investigation to another body, there could be a conflict between the privacy of individuals and the expectations for transparency for an investigation.

There will be a question of priority for an investigation; at what point does the search for "truth" and transparency outweigh privacy concerns or bystanders, victims, or people living in the area. The risk of harm to individuals will be context specific but must be considered in any investigation.

Findings from attribution, or even the existence of an investigation, can impact the lives of a population or individuals. Within a community, there could be blowback against individuals who cooperate with the investigation. An entire population could be discriminated against based on results of an investigation. If an investigation looks into one group over the course of an investigation but later clears that group, the group could be subject to mis/disinformation campaigns. An investigation team will need to be cognizant of these risks and work to minimize potential harm.

Logistics

All parts of in investigation will be heavily impacted by logistics. Logistics will include accessibility and availability of samples, laboratories, supplies, and investigators, access to areas of interest and victims, laboratory capacities, transport of samples and investigators, and coordination between stakeholders. Any type of WMD investigation includes logistical challenges, but a biological event investigation will be especially challenging given the time constraints for collecting and processing biological samples. Time constraints will put immense pressure on an investigatory team to overcome logistical challenges. A support team

focusing on logistics would be helpful for investigators.

In addition to the time considerations, a particular challenge for a multilateral investigation will be coordination between countries, relevant international bodies such as UN agencies, and the investigators. In the case of an UNSGM investigation, there will be a centralized investigatory team, which will lower the burden of having to coordinate between investigatory teams. However, if there is not a centralized team, then there will be another layer of logistical challenge to coordinate between each unique team. In the case of the UNSGM system, the investigatory teams will have to coordinate communications between laboratories and from laboratories to the investigatory team. Given the complexities of biological investigations, there will be several players in any scenario, each of which will add to logistical complexity.

Politics

Geopolitics will heavily impact if, how, when, and by whom a biological event is investigated. Countries with strained international relations may not report suspicious cases or events to other countries or international bodies. Countries that do not trust or want international involvement may decline to participate in an investigation with other countries or deny entry to international organizations, like WHO. Some countries may have sovereignty or national security concerns, which could dictate who they are willing to share evidence and information with. Competing political agendas will also influence accessibility to data and evidence. Countries could restrict visas or use sanctions to affect an investigatory team composition or intimidate a team in hopes of influencing final findings. Politics is a driver of mis- and disinformation and could be employed by political operatives to control a narrative around a biological event. While the geopolitical landscape can change, there are several areas of planning that could be undertaken before an event. Bilateral and multilateral trainings between allies can help set a foundation for future investigations.

5.3.2 Attribution Capabilities

Two overarching capabilities were identified for biological attribution: scientific capabilities and investigatory capabilities. Scientific capabilities refer to technology and scientific knowledge for attribution while the investigatory capabilities refer to all other capabilities needed to conduct the investigation for attribution.

Scientific Capabilities

Scientific capabilities include technology available for analysis, sample collection, or processing. There are many ways to improve scientific capabilities for attribution. Advances in sequencing capabilities or algorithms for attribution would fall under this category. Improved sensitivity and specificity of for tests or improved yield would increase the scientific capabilities for attribution. Increased samples available in databases or improved cataloguing within databases can also increase this area of capability for attribution. Availability of appropriate controls or validated protocols falls here as well.

Investigatory Capabilities

Investigatory capabilities include workforce availability, laboratory infrastructure, a support structures such as administrative and logistical support. Availability of individuals with training in investigations and the technical requirements is vital for any investigation but not a given. The availability of laboratories with the appropriate tools and regents to run required tests will similarly affect the ability of the investigation to move forward. Other notable factors for investigatory capabilities include available budget and secure telecommunications resources.

5.3.3 Investigatory Strategy

Several components were identified for the investigatory strategy. The components can span all or parts of the investigatory process, as illustrated in Figure X. Each of these components of the strategy will be key considerations for implementing an investigation for attribution. These will be major decision points for an investigatory team, where prior planning can provide insight into different options, but a final decision will be dependent on the event itself and can change from one event to another.

Investigatory Team Composition

The composition of the investigatory team will be a critical factor for any investigation but will need to be unique to the situation. The type of agent will determine if a bacteriologist, mycologist, toxicologist, or virologist is better suited for the team. Where the event occurred will impact the nationalities of the team members. Ensuring all team members have appropriate training for not just the technical considerations

of the investigation but also skills such as public speaking or familiarity with relevant jurisdictions will be important. Language barriers will need to be considered and addressed. Ensuring all necessary areas of expertise will be critical for the direction and legitimacy of the investigation.

Time Since Event

The time it takes to recognize there is an abnormal health condition or epidemiological pattern will heavily impact the strategy for an investigation. If too much time has passed, there may not be viable samples at the origin site. Time since event directly impacts sample availability and degradation, which will likely be the foundation of an investigation. Time since the event could also impact public perception of the investigation, which could lead to challenges with misinformation or obtaining statements from witnesses.

Epidemiological Context

Historical rates of disease, ongoing outbreaks, prevalence data will be important for identifying if an event has occurred and will be important evidence to demonstrate if the event is naturally occurring or not. Epidemiological data will be critical evidence for assessing the nature of the event and contextualizing results. Epidemiological data could include sensitivity and specificity of test of diagnostic tests, rates of use of test (to demonstrate how easily could the agent be spreading and missed by health care system/public health), novelty of agent, demographics of affected people, historical outbreak patterns, exposure history, zoonotic considerations like presence of known host or intermediate species in area or reservoir species, and ecological changes happening before or at time of event that could impact disease patterns.

Training

Laboratories and individuals that could be utilized in an investigation will need training to ensure they are familiar with protocols and standards. There will likely need to be a balance between technical expertise and level of training for investigations. The leading expert in a given pathogen may not be the best pick for an investigatory team if they have no training or familiarity with evidence collection standards or working with international organizations. There is a wide breadth of trainings that could be relevant for attribution. Training could include formal education in a technical area as well as certifications in a skill set or participation

in exercises related to attribution.

Sample Quality, Handling, Storage, and Transport

A key factor impacting evidence for attribution will be the availability and quality of samples. All downstream efforts will hinge on the samples collected. Type of sample (urine, CSF, blood, sputum, etc.), environmental sample availability, and timing of sample collection directly relates to evidence quality and therefor investigation quality. Ensuring samples are handled properly and follow all relevant protocols, especially related to chain of evidence, will impact validity of the investigation. Storage of samples for re-analysis or quality controls will be important for trust in the investigation. All aspects of handling samples can impact results generated from those samples.

Forensics

Forensic analysis of evidence will be the central feature of any investigation. The strategy for forensics will be highly dependent on the type and nature of samples collected. While different weights may be given to different types of evidence, there must be capabilities to collect, analyze, and interpret all different types of evidence that exists for a given event. There are three categories of forensics that will be relevant for biological attribution.

Biological Agent Forensics: This category refers to forensics of the agent(s) at the center of the event. The agent could be a bacterium, virus, fungi, or toxin produced by a biological agent. Forensics on these agents will require different protocols and safety measures than most criminal investigations require. Specialized infrastructure like specialized laboratory space built for containment or elevated biosafety strategies may be required. The types of analysis will be unique for these samples as well and may include assessing phylogenetic models, molecular clock estimates, selection pressure and genetic drift analysis, ability to culture the agent, mutational signatures, mobile genetic elements (plasmids, phages, resistance genes), and markers of engineering, purity, or origin within an organism.

Conventional Biological Sample Forensics: Conventional biological samples include biological samples that are regularly and routinely analyzed in most forensic laboratories and therefore there are many facilities and experts available to handle this type of evidence. This category includes analyzing trace DNA, bodily

fluids, fingerprints, testing biological samples for noninfectious agents that could be causing disease like heavy metals or other toxic chemical exposures, examine autoimmune markers of affected individuals, rule out genetic disorders, or compare medical records to look for consistency or unique clinical markers amongst victims.

Conventional Non-Biologic Sample Forensics: This category refers to analysis of evidence that is not biologic in nature. For example, CCTV, witness statements, supply chain analysis, human and signal intelligence sources, geospatical mapping, delivery system analysis, potential actor capability and motivation assessments (including technical capabilities, history of threats, political or ideological motivations), analyzing physical remnants (e.g., dispersal devices, bomb fragments), analyzing lab access, dual use technology assessment and availability analysis, or evaluate potential for false flag operations or framing are all conventional non-biologic sample forensics.

Data Interpretability

The results from analyzing samples and interpretation and reporting of such results heavily relies on the interpretability of the data. A good investigatory strategy will consider data interpretability when choosing methods for analysis of sample. Data interpretability will also encompass how results are construed into the overall findings and subsequently reported. Incorrect or overstated interpretations of one piece of evidence could lead to the entire investigation being discounted. Statistical methods utilized, baseline data availability, use of models and their associated parameters, and statements relaying confidence measures are all part of data interpretability. Deciding how much contextualization or interpretation of results, compared to just reporting raw results, will be a key challenge for the investigatory team. The team will have to determine the line between what is included and what is not included in final reports. Determining which hypotheses are addressed and tested, if alternative scenarios are assessed, and narratives about evidence convergence will be crucial decisions.

Media Strategy

Understanding that many people will not read the report directly and will instead hear results through media sources will be critical for the investigation. There must be a media strategy in place for reporting results.

Questions such as how will results be published and to whom, if different versions of the final report will be necessary, at which points in an investigation will an update be publicly provided, and strategies for combating mis/disinformation will be important to consider throughout an investigation. Implications of how, when, and where results are reported could have severe consequences for those impacted by the event. Reporting of results could lead to retaliation by suspects or deter would-be actors. The investigatory team will need to consider how much evidence is enough to share to support their findings while protecting privacy of those affected or ensuring dangerous information is not irresponsibly disseminated. **A media strategy will need to include plans for monitoring social media and other websites to monitor for conspiracy theories or misuse of information from the investigation. Plans should be built for addressing significant theories that persist or rapidly amplify, especially those that appear to be pushed by bots.** Results from an investigation may impact response to an event that is ongoing and determining the right time to share that information could save lives.

5.3.4 Relationships Between the Components of the Framework

Relationships between elements in the model are labeled alphanumerically, with C1-3 representing relationships between contextual factors, attribution capabilities, and the investigatory strategy. Relationships between elements within the investigatory strategy category are represented S1-13.

Relationship C1 (between contextual factors and attribution capabilities) is bidirectional. Contextual factors will drive capabilities in several ways. Political factors influence science through funding and priorities for research, which directly impacts the scientific capabilities for attribution. All contextual factors can impact the investigatory capabilities; politics will determine budget, logistical factors will influence who can investigate and when, policy factors will impact which tools are available to investigators. However, knowledge of the scientific and investigatory capabilities can impact the contextual factors as well. Attribution itself is considered a deterrent for biological weapons use, which impacts the political context. Understanding capabilities can drive policy improvements or changes. Capabilities can improve, causing contextual factors to change, or contextual factor changes could directly or indirectly impact capabilities.

Relationships C2 and C3 represent the connection between contextual factors to investigation strategy and attribution capabilities to investigatory strategy, respectively. The strategy for conducting the investi-

gation will be heavily influenced by the contextual factors and the available capabilities. Major changes to contextual factors or capabilities over the course of the investigation could require the investigatory plan to be overhauled, which would make the endeavor vulnerable.

Contextual factors will influence strategy for the investigation in several ways. Policy, including international norms and jurisdictional regulations, will dictate how an investigation is designed, its audience, and the scope of the investigation. Ethical considerations will impact how the investigation is designed, who is involved, and how the results are reported. Logistical constraints will have to be factored into the design of the investigation and could impact strategy determinations, especially related to the initiation of the investigation and evidence collection. Finally, political considerations will heavily dictate how the investigation moves forward and how it is reported.

Attribution capabilities will determine the limits of an investigation and must be factored into the strategy. Scientific capabilities will determine what is possible to determine from the evidence and, in some cases, what evidence it is possible to obtain. Investigatory capabilities will determine the resources available to make the investigation possible. A strategy must consider both capabilities when designing the investigation as both areas of capability impact all components of a strategy.

The elements of the investigatory strategy will influence one another. The investigatory team should be comprised of people with the relevant expertise and sufficient training for an investigation (S1). The epidemiological context will impact the composition of the investigatory team (S2) as different epidemiological contexts will determine the types of expertise needed to thoroughly understand the nature of an event and assess potential trends in disease across a population as part of a forensic analysis (S8). The time since the event occurred will impact how much information about the epidemiological context is available (S3) and the possible evidence available for forensic analysis (S4). The investigatory team will determine which forensic methods are chosen and how results are interpreted, so initial understandings of what forensics methodologies will be needed should be considered when creating the investigatory team (S5). The investigatory team should be familiar enough with both technical considerations for sampling and investigatory needs to be ensure sample integrity and quality through all stages of a sample's life cycle (S6), recognizing that the sample quality and how the sample is handled will directly influence the forensics analysis on the sample (S7). Quality of samples will directly impact how readily the results can be interpreted (S9). The

forensic methods chosen will also determine interpretability of the data and assessing interpretability needs can direct forensic methodologies chosen (S10). Findings from forensic analyses (S11) and the broader context within which individual findings are interpreted (S12) will impact how the data is shared to the public via the media. Recognizing the importance of communicating with the media, the investigatory team should include people trained and comfortable relaying highly technical or contentious data with the media (S13).

5.4 Conclusion

This framework was developed to better conceptualize biological attribution. In doing so, areas of significant impact for attribution and key decision-making points are identified. Such knowledge can aid in identifying priority areas for planning, training, or research to strengthen attribution. This framework remains theoretical. Future work should challenge the model through simulations, tabletop exercises, or real investigations.

Contextual factors and capabilities will determine the limitations of an investigation. They will define the broad outline of an investigation while the investigatory strategy determines the details. Major changes in contextual factors or capabilities will fundamentally alter the environment in which the investigation is conducted. As such, the investigatory strategy must be flexible and resilient to withstand such changes. Planning for attribution should heavily focus on contextual factors and capabilities as they will provide the foundation for an investigation. The investigatory strategy will be different for every event and therefore difficult to sufficiently plan. However, contextual factors and capabilities can be targeted well before an event and by strengthening capabilities and understanding contextual factors, appropriate investigation strategies can be built better.

To further strengthen attribution abilities, this framework could be used to develop a model to enable formal assessment for attribution planning and readiness. Formal assessments for attribution can help develop strategies to aid funders and other stakeholders in determining where resources should be allocated to maximize readiness for an biological event.

Chapter 6

Conclusion

6.1 Summary of Findings

A mixed methods approach was used in this dissertation to examine how different stakeholders interpret and weigh evidence for attribution, exploring how investigatory processes might be perceived as credible and actionable, and developing a conceptual framework to guide future attribution efforts.

Chapter 3 included interviews with 41 experts from a range of disciplines showed little consensus on what constitutes "attribution." While some viewed it as determining whether an event was natural, accidental, or deliberate, others defined it narrowly as assigning culpability for intentional misuse. Factors impacting the experts' views on attribution were identified. Trust, legitimacy, and international cooperation were recurring concerns, alongside recognition of technical and organizational constraints. Findings show that credibility depends less on the technical rigor of evidence alone and more on whether the process is transparent, inclusive, and insulated from politicization. Beliefs that are strongly held and considered obvious to someone with one background sometimes were directly contradictory to beliefs held by others with different backgrounds, such as on the question of whether an investigatory team should use standard methods exclusively or also use more cutting-edge technology. Results from this research demonstrate why effort must be put into developing consensus and clarity on protocols for attribution before an event happens.

Chapter 4 assessed variability in evidence interpretation. Scenario-based exercises revealed striking differences in how participants weighed evidence. Genetic data was often perceived as highly probative,

yet interpretations diverged widely, even among experts. Some relied heavily on sequence analysis, while others discounted it entirely. Epidemiological, contextual, and intelligence evidence, though often seen as less definitive, significantly shaped participants' judgments. Overall, the weight and interpretation of evidence was different between specialties. All subspecialities expected a convergence of multiple streams of evidence to perceive the investigation as credible.

Chapter 5 developed an integrated conceptual framework. The framework developed in Chapter 5 maps three domain: contextual factors (geopolitics, institutions, resources), capabilities (technical and organizational), and investigatory strategy (evidence handling, reporting, communication). The framework illustrates how technical advances, while critical, are insufficient on their own; the credibility and actionability of attribution depend equally on governance, communication, and social trust. The framework also provides a roadmap for planning and capacity-building before a biological event occurs. Actionability depends on clear governance, communication strategies that manage uncertainty, and international trust. Without these, even strong scientific evidence may fail to achieve legitimacy. While technical advances are critical for feasibility of attribution, the context of the investigation is more likely to cause the endeavor to fail.

6.2 Implications for Future Research, Policy, and Practice

The findings of this dissertation underscore that biological attribution is not a purely scientific or technical process, but a multidisciplinary effort shaped by politics, law, communication, and ethics. The implications extend across research, policy, and practice, pointing to several areas where further work is needed to strengthen readiness for future biological events.

6.2.1 Future Research

Research should continue to probe how different stakeholders interpret, prioritize, and act upon evidence. The scenario exercises in this study revealed wide variation in how even experts use genetic, epidemiological, and intelligence data. Future work should examine how interdisciplinary teams negotiate these differences in real-time settings, such as simulations or tabletop exercises, to better capture the dynamics of group decision making. Research should also expand beyond technical experts to include perspectives from legal professionals and community representatives, especially those less familiar with biosecurity, as

attribution processes will ultimately be judged by diverse audiences. Tabletop exercises with elected officials would also be key, as elected politicians are most likely going to be the people making decisions about what to do with results from an investigation. Understanding how comfortable elected politicians are with different types of reporting, such as a report heavy with statistics versus a summary report that lacks granular detail but is written for a non-scientific audience, will be helpful for future investigatory teams. Additionally, understanding how elected officials expect to be use their own staff and advisors to approach results will be important.

Another priority is the evaluation of emerging technologies. High-throughput sequencing, -omics platforms, and machine learning tools hold promise for attribution, but questions remain about their reproducibility, interpretability, and vulnerability to manipulation. Systematic studies should assess not only technical performance but also how these tools are perceived in judicial, diplomatic, and public arenas. Finally, research should explore ethical frameworks for attribution, particularly mechanisms to balance transparency with protections against stigmatization and premature conclusions.

6.2.2 Policy Implications

At the policy level, the findings highlight the urgent need for clear governance structures that delineate roles and responsibilities across public health, law enforcement, intelligence, and international organizations. Attribution efforts risk delay or conflict without established frameworks for coordination. Policymakers should prioritize the development of protocols and legal agreements that enable rapid, cooperative investigations while safeguarding due process.

Internationally, attribution remains hampered by limited trust and voluntary cooperation. Strengthening global capacity will require not only investment in laboratories and expertise but also mechanisms for data sharing, joint investigations, and dispute resolution. Policymakers should consider expanding the mandate of existing mechanisms, such as the United Nations Secretary-General's Mechanism, or developing new institutions specifically focused on biological attribution. Establishing international standards for evidence collection, chain of custody, and data interpretation would enhance credibility and comparability across investigations.

Finally, policies must anticipate the communication challenges that accompany biological attribution.

Misinformation and politicization can undermine even the most rigorous investigation. Developing strategies for transparent, timely, and audience-appropriate communication should be a core policy priority, supported by training for investigators and spokespeople alike.

6.2.3 Implications for Practice

For practitioners, the findings emphasize the importance of capacity-building and preparedness before a crisis occurs. Exercises and simulations should be used to test investigatory protocols, practice interdisciplinary collaboration, and identify gaps in resources and expertise. Building relationships across agencies and nations in advance will be critical for ensuring cooperation during high-pressure investigations.

Practitioners must also prepare to engage with the public and media. The credibility of attribution efforts depends not only on technical accuracy but also on the perceived legitimacy of the process. Training investigators to communicate uncertainty responsibly and to manage expectations will be essential for sustaining trust.

Taken together, these implications suggest that strengthening biological attribution will require an approach that integrates technical innovation, institutional reform, and public engagement. Future research can provide the evidence base, policy can establish the structures and standards, and practice can operationalize these insights in real world investigations.

6.3 Final Reflections

This project was conceptualized during the early months of the COVID-19 pandemic. The world changed in a matter of months and people were scared. Within a month of the WHO being notified of the first cases, there were already conspiracy theories circulating online about where the virus had come from. The conspiracy theories quickly made it to the mouths of politicians, who were also scrambling to figure out how to respond to the threats posed by a pandemic. Even five years later, the origins of COVID-19 are being debated and used to justify policy. The 2025 "report" from the White House demonstrates why it is critical to study origins of pandemics thoroughly and within trusted and well-developed systems (115). The origin of a biological event can easily be politicized, leading to ripple effects that lead to decreased trust in science, medicine, and public health. When this mistrust is allowed to grow through a wide enough population, the

government is able to make decisions that are counter to decades of evidence and harm the people they are supposed to be serving (182). If COVID-19 origins had not been so ambiguous, there might not have been as large of an anti-vaccination campaign. Without the anti-COVID vaccine push, anti-vaccinators might not be as powerful of an entity today.

I believe that if there had been a more transparent and thorough investigation for COVID-19 origins, there wouldn't have been as big an opportunity for conspiracy theorists to reach the mainstream. While most scientists with training in viral genetics and evolution agree on the origins, their word was not enough for scientists and others who did not understand the limits of their own knowledge. They didn't know what they didn't know. They were experts in their own field, but not in the highly specialized field of viral genetics and evolution. Instead, many of these people were left to use circumstantial evidence. The lack of transparency from China, the first cases being in the same providence as a virology lab working on coronaviruses, and missing data combined to make a compelling story that the virus could have come from a laboratory. When the more technical genetic evidence is ignored or misunderstood, then the story of a laboratory escape seems plausible.

Because there were people trusted in fields related to virology claiming the lab escape hypothesis was the most likely answer, it paved the way for conspiracy theories to grab a greater foothold in public health areas. Trust eroded in public health as increased doubt was cast on public health agencies and practitioners, with COVID-19 origins being one of many areas where the public lost confidence. Now, in the US, we have seen defunding of public health agencies and declining access to COVID-19 vaccines (183; 184). The kneecapped public health system and defunded biomedical science research is a threat to everyone's health.

I don't think the debacle of COVID-19 origins is solely responsible for the current state of public health in this country, but I do think it played a part in allowing us to get to this point. This is a prime example of why it is so important to continue to research and strengthen attribution capabilities. Building stronger implementation systems and planning ahead for how to communicate about origins, investigations, and attribution, will be critical for repairing trust in our public health institutions.

Bibliography

(1) Gopal Das, Shailendra Pratap Jain, Durairaj Maheswaran, Rebecca J. Slotegraaf, and Raji Srinivasan. 2021. Pandemics and marketing: insights, impacts, and research opportunities. *Journal of the Academy of Marketing Science*, 49(5):835–854. URL: https://link.springer.com/10.1007/s11747-021-00786-y.

(2) William Mo, Christopher A. Vaiana, and Chris J. Myers. 2024. The need for adaptability in detection, characterization, and attribution of biosecurity threats. *Nature Communications*, 15(1):10699. URL: https://www.nature.com/articles/s41467-024-55436-y.

(3) A Adel. A Conceptual Framework to Improve Cyber Forensic Administration in Industry 5.0: Qualitative Study Approach. *Forensic Sciences*, 2(1):111–129. URL: https://www.mdpi.com/2673-6756/2/1/9.

(4) F M G Wong, T Hinton, and D K Smith. Overview of Nuclear Forensics in Support of Investigations. Technical Report IAEA-CN-218-48, IAEA.

(5) S. Mishra, S. Anilkumar, and A. Vinod Kumar. 2023. Nuclear Forensics: Role of Radiation Metrology. In *Handbook of Metrology and Applications*, pages 2293–2319. Springer, Singapore. URL: https://link.springer.com/rwe/10.1007/978-981-99-2074-7_134.

(6) IAEA. 2015. *Nuclear Forensics in Support of Investigations: Implementing Guide*. Number v.2-G (Rev. 1) in IAEA Nuclear Security Series No. 2-G (Rev. 1). IAEA, Vienna.

(7) Sukanta Maity, Amar Pant, Sandeep Police, and Amit Verma. 2024. Role of Nuclear Forensics in Preparedness and Response in Radiation Emergencies. In Dinesh Kumar Aswal, editor, *Handbook*

on Radiation Environment, Volume 1: Sources, Applications and Policies, pages 581–616. Springer Nature, Singapore. URL: https://doi.org/10.1007/978-981-97-2795-7_19.

(8) Yasuo Seto. 2015. Chapter 60 - On-Site Detection of Chemical Warfare Agents. In Ramesh C. Gupta, editor, *Handbook of Toxicology of Chemical Warfare Agents (Second Edition)*, pages 897–914. Academic Press, Boston. URL: https://www.sciencedirect.com/science/article/pii/B9780128001592000609.

(9) Andrew Clements, Ian Mendenhall, and Daniel Schar. 2024. Understanding How and Where Pathogens Emerge: Preparedness and Response for Zoonotic Diseases. In *Principles and Practice of Emergency Research Response [Internet]*. Springer. URL: https://www.ncbi.nlm.nih.gov/books/NBK613994/.

(10) James M. Hughes, Mary E. Wilson, Brian L. Pike, Karen E. Saylors, Joseph N. Fair, Matthew LeBreton, Ubald Tamoufe, Cyrille F. Djoko, Anne W. Rimoin, and Nathan D. Wolfe. 2010. The Origin and Prevention of Pandemics. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*, 50(12):1636–1640. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2874076/.

(11) Raina K. Plowright, Aliyu N. Ahmed, Tim Coulson, Thomas W. Crowther, Imran Ejotre, Christina L. Faust, Winifred F. Frick, Peter J. Hudson, Tigga Kingston, P. O. Nameer, M. Teague O'Mara, Alison J. Peel, Hugh Possingham, Orly Razgour, DeeAnn M. Reeder, Manuel Ruiz-Aravena, Nancy B. Simmons, Prashanth N. Srinivas, Gary M. Tabor, Iroro Tanshi, Ian G. Thompson, Abi T. Vanak, Neil M. Vora, Charley E. Willison, and Annika T. H. Keeley. 2024. Ecological countermeasures to prevent pathogen spillover and subsequent pandemics. *Nature Communications*, 15(1):2577. Publisher: Nature Publishing Group. URL: https://www.nature.com/articles/s41467-024-46151-9.

(12) Vishakha Vashisht, Ashutosh Vashisht, Ashis K. Mondal, Jaspreet Farmaha, Ahmet Alptekin, Harmanpreet Singh, Pankaj Ahluwalia, Anaka Srinivas, and Ravindra Kolhe. 2023. Genomics for Emerging Pathogen Identification and Monitoring: Prospects and Obstacles. *BioMedInformat-*

ics, 3(4):1145–1177. Publisher: Multidisciplinary Digital Publishing Institute. URL: https://www.mdpi.com/2673-7426/3/4/69.

(13) Gregory Lewis, Jacob L. Jordan, David A. Relman, Gregory D. Koblentz, Jade Leung, Allan Dafoe, Cassidy Nelson, Gerald L. Epstein, Rebecca Katz, Michael Montague, Ethan C. Alley, Claire Marie Filone, Stephen Luby, George M. Church, Piers Millett, Kevin M. Esvelt, Elizabeth E. Cameron, and Thomas V. Inglesby. 2020. The biosecurity benefits of genetic engineering attribution. *Nature Communications*, 11(1):6294. URL: https://www.nature.com/articles/s41467-020-19149-2.

(14) Ingrid Wiechmann, Michaela Harbeck, and Gisela Grupe. 2010. Yersinia pestis DNA Sequences in Late Medieval Skeletal Finds, Bavaria. *Emerging Infectious Diseases*, 16(11):1806–1807. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3294523/.

(15) Giovanna Morelli, Yajun Song, Camila J. Mazzoni, Mark Eppinger, Philippe Roumagnac, David M. Wagner, Mirjam Feldkamp, Barica Kusecek, Amy J. Vogler, Yanjun Li, Yujun Cui, Nicholas R. Thomson, Thibaut Jombart, Raphael Leblois, Peter Lichtner, Lila Rahalison, Jeannine M. Petersen, Francois Balloux, Paul Keim, Thierry Wirth, Jacques Ravel, Ruifu Yang, Elisabeth Carniel, and Mark Achtman. 2010. Yersinia pestis genome sequencing identifies patterns of global phylogenetic diversity. *Nature Genetics*, 42(12):1140–1143. Publisher: Nature Publishing Group. URL: https://www.nature.com/articles/ng.705.

(16) Maria A. Spyrou, Lyazzat Musralina, Guido A. Gnecchi Ruscone, Arthur Kocher, Pier-Giorgio Borbone, Valeri I. Khartanovich, Alexandra Buzhilova, Leyla Djansugurova, Kirsten I. Bos, Denise Kühnert, Wolfgang Haak, Philip Slavin, and Johannes Krause. 2022. The source of the Black Death in fourteenth-century central Eurasia. *Nature*, 606(7915):718–724. Publisher: Nature Publishing Group. URL: https://www.nature.com/articles/s41586-022-04800-3.

(17) John Kelly. 2006. *The Great Mortality: An Intimate History of the Black Death, the Most Devastating Plague of All Time*. Harper Perennial, New York City.

(18) Howard Brody, Michael Russell Rip, Peter Vinten-Johansen, Nigel Paneth, and Stephen Rachman.

2000. Map-making and myth-making in Broad Street: the London cholera epidemic, 1854. *The Lancet*, 356(9223):64–68. URL: https://linkinghub.elsevier.com/retrieve/pii/S0140673600024429.

(19) Nigel Stephen Walford. 2020. Demographic and social context of deaths during the 1854 cholera outbreak in Soho, London: a reappraisal of Dr John Snow's investigation. *Health & Place*, 65:102402.

(20) Michael Worobey, Jim Cox, and Douglas Gill. 2019. The origins of the great pandemic. *Evolution, Medicine, and Public Health*, 2019(1):18–25. URL: https://doi.org/10.1093/emph/eoz001.

(21) Michaela E. Nickol and Jason Kindrachuk. 2019. A year of terror and a century of reflection: perspectives on the great influenza pandemic of 1918–1919. *BMC Infectious Diseases*, 19(1):117. URL: https://doi.org/10.1186/s12879-019-3750-8.

(22) Jeffery K. Taubenberger, David Baltimore, Peter C. Doherty, Howard Markel, David M. Morens, Robert G. Webster, and Ian A. Wilson. 2012. Reconstruction of the 1918 Influenza Virus: Unexpected Rewards from the Past. *mBio*, 3(5):10.1128/mbio.00201–12. Publisher: American Society for Microbiology. URL: https://journals.asm.org/doi/10.1128/mbio.00201-12.

(23) Trevor Hoppe. 2018. "Spanish Flu": When Infectious Disease Names Blur Origins and Stigmatize Those Infected. *American Journal of Public Health*, 108(11):1462–1464. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6187801/.

(24) Michael Worobey, Thomas D. Watts, Richard A. McKay, Marc A. Suchard, Timothy Granade, Dirk E. Teuwen, Beryl A. Koblin, Walid Heneine, Philippe Lemey, and Harold W. Jaffe. 2016. 1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America. *Nature*, 539(7627):98–101.

(25) T. Baumgartl, M. Petzold, M. Wunderlich, M. Hohn, D. Archambault, M. Lieser, A. Dalpke, S. Scheithauer, M. Marschollek, V. M. Eichel, N. T. Mutters, Highmed Consortium, and T. Von Landesberger. 2021. In Search of Patient Zero: Visual Analytics of Pathogen Transmission Pathways in Hospitals. *IEEE transactions on visualization and computer graphics*, 27(2):711–721.

(26) Maohui Feng, Qiong Ling, Jun Xiong, Anne Manyande, Weiguo Xu, and Boqi Xiang. 2021. Occupational Characteristics and Management Measures of Sporadic COVID-19 Outbreaks From June 2020 to January 2021 in China: The Importance of Tracking Down "Patient Zero". *Frontiers in Public Health*, 9:670669. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8119752/.

(27) Richard A. McKay. 2017. *Patient Zero and the Making of the AIDS Epidemic*. University of Chicago Press, Chicago, IL. URL: https://press.uchicago.edu/ucp/books/book/chicago/P/bo16463356.html.

(28) Nino Antulov-Fantulin, Alen Lančić, Tomislav Šmuc, Hrvoje Štefančić, and Mile Šikić. 2015. Identification of Patient Zero in Static and Temporal Networks: Robustness and Limitations. *Physical Review Letters*, 114(24):248701. Publisher: American Physical Society. URL: https://link.aps.org/doi/10.1103/PhysRevLett.114.248701.

(29) Colin J. Worby, Philip D. O'Neill, Theodore Kypraios, Julie V. Robotham, Daniela De Angelis, Edward J. P. Cartwright, Sharon J. Peacock, and Ben S. Cooper. 2016. Reconstructing transmission trees for communicable diseases using densely sampled genetic data. *The annals of applied statistics*, 10(1):395–417. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4817375/.

(30) Ralf Reintjes and Aryna Zanuzdana. 2009. Outbreak Investigations. In *Modern Infectious Disease Epidemiology*, pages 159–176. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7187955/.

(31) Andreas G. Nerlich, Bettina Schraut, Sabine Dittrich, Thomas Jelinek, and Albert R. Zink. 2008. Plasmodium falciparum in Ancient Egypt. *Emerging Infectious Diseases*, 14(8):1317–1319. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC2600410/.

(32) Weimin Liu, Yingying Li, Gerald H. Learn, Rebecca S. Rudicell, Joel D. Robertson, Brandon F. Keele, Jean-Bosco N. Ndjango, Crickette M. Sanz, David B. Morgan, Sabrina Locatelli, Mary K. Gonder, Philip J. Kranzusch, Peter D. Walsh, Eric Delaporte, Eitel Mpoudi-Ngole, Alexander V.

Georgiev, Martin N. Muller, George M. Shaw, Martine Peeters, Paul M. Sharp, Julian C. Rayner, and Beatrice H. Hahn. 2010. Origin of the human malaria parasite Plasmodium falciparum in gorillas. *Nature*, 467(7314):420–425. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC2997044/.

(33) Dorothy E. Loy, Weimin Liu, Yingying Li, Gerald H. Learn, Lindsey J. Plenderleith, Sesh A. Sundararaman, Paul M. Sharp, and Beatrice H. Hahn. 2017. Out of Africa: origins and evolution of the human malaria parasites *Plasmodium falciparum* and *Plasmodium vivax*. *International Journal for Parasitology*, 47(2):87–97. URL: https://www.sciencedirect.com/science/article/pii/S0020751916301229.

(34) Barbara Herwaldt. 2001. Laboratory-acquired parasitic infections from accidental exposures. *Clinical Microbiology Reviews*, 14(4):659–688, table of contents.

(35) Stuart D. Blacksell, Sandhya Dhawan, Marina Kusumoto, Khanh K. Le, Kathrin Summermatter, Joseph O'Keefe, Joseph P. Kozlovac, Salama S. Almuhairi, Indrawati Sendow, Christina M. Scheel, Anthony Ahumibe, Zibusiso M. Masuku, Allan M. Bennett, Kazunobu Kojima, David R. Harper, and Keith Hamilton. 2024. Laboratory-acquired infections and pathogen escapes worldwide between 2000 and 2021: a scoping review. *The Lancet Microbe*, 5(2):e194–e202. Publisher: Elsevier. URL: https://www.thelancet.com/journals/lanmic/article/PIIS2666-5247(23)00319-1/fulltext.

(36) Z. F. Dembek, M. G. Kortepeter, and J. A. Pavlin. 2007. Discernment between deliberate and natural infectious disease outbreaks. *Epidemiology and Infection*, 135(3):353–371.

(37) Filippa Lentzos, Gregory Koblentz, and Joseph Rodgers. 2022. The Urgent Need for an Overhaul of Global Biorisk Management. *CTC Sentinel*, 15(4). URL: https://ctc.westpoint.edu/the-urgent-need-for-an-overhaul-of-global-biorisk-management/.

(38) Committee on Review of the Scientific Approaches Used During the FBI's Investigation of the 2001 Bacillus Anthracis Mailings and National Research Council. 2011. *Review of the Scientific Ap-*

proaches Used during the FBI's Investigation of the 2001 Anthrax Letters. National Academies Press (US), Washington (DC). URL: http://www.ncbi.nlm.nih.gov/books/NBK209412/.

(39) James Weaver. 1985. The Town That Was Poisoned. URL: https://en.wikisource.org/wiki/The_Town_That_Was_Poisoned/Congressman_Weaver.

(40) Tsuneishi Keiichi. 2005. Unit 731 and the Japanese Imperial Army's Biological Warfare Program. *Asia-Pacific Journal*, 3(11):e24. Publisher: Cambridge University Press & Assessment. URL: https://www.cambridge.org/core/journals/asia-pacific-journal/article/unit-731-and-the-japanese-imperial-armys-biological-warfare-program/815CBD55CEB1CFCBA06034704ECAFC1E.

(41) Kishor Johnson. 2022. A Scientific Method to the Madness of Unit 731's Human Experimentation and Biological Warfare Program. *Journal of the History of Medicine and Allied Sciences*, 77(1):24–47.

(42) Ronald A. Greenfield, Brent R. Brown, James B. Hutchins, John J. Iandolo, Rhett Jackson, Leonard N. Slater, and Michael S. Bronze. 2002. Microbiological, biological, and chemical weapons of warfare and terrorism. *The American Journal of the Medical Sciences*, 323(6):326–340.

(43) Rashed Alremeithi, Natalie Sullivan, Hannah Checkeye, Maryann Mazer-Amirshahi, and Ali Pourmand. 2023. A clinical approach to an unidentified aerosolized bioterrorism agent: a narrative review for emergency providers. *Clinical and Experimental Emergency Medicine*, 10(2):147–156. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10350351/.

(44) P. Keim, K. L. Smith, C. Keys, H. Takahashi, T. Kurata, and A. Kaufmann. 2001. Molecular investigation of the Aum Shinrikyo anthrax release in Kameido, Japan. *Journal of Clinical Microbiology*, 39(12):4566–4567.

(45) M. Papaloucas, C. Papaloucas, and A. Stergioulas. 0. Ricin and the Assassination of Georgi Markov. *Pakistan Journal of Biological Sciences*, 11(19):2370–2371. URL: https://scialert.net/fulltext/?doi=pjbs.2008.2370.2371.

(46) Seth Carus. 2001. Bioterrorism and Biocrimes: The Illicit Use of Biological Agents Since 1900:. Technical report, National Defense University, Fort Belvoir, VA. URL: https://apps.dtic.mil/sti/citations/tr/ADA402108.

(47) Bradley D Stein, Terri L Tanielian, David P Eisenman, Donna J Keyser, M Audrey Burnam, and Harold A Pincus. 2004. Emotional and Behavioral Consequences of Bioterrorism: Planning a Public Health Response. *The Milbank Quarterly*, 82(3):413–455. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2690224/.

(48) Marc Ostfield. 2004. Historical Uses of Biological Agents as Weapons. *Sais Review*, 24:135–137.

(49) V. Barras and G. Greub. 2014. History of biological warfare and bioterrorism. *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, 20(6):497–502.

(50) W. Seth Carus. 2015. The History of Biological Weapons Use: What We Know and What We Don't. *Health Security*, 13(4):219–255. Publisher: Mary Ann Liebert, Inc., publishers. URL: https://www.liebertpub.com/doi/10.1089/hs.2014.0092.

(51) Stephen Papagiotas and Kelly Shannon. 2018. Suspected Intentional Use Of Biologic And Agents. In *The CDC Field Epidemiology Manual*. Centers for Disease Control and Prevention.

(52) Gregory D. Koblentz and Brian M. Mazanec. 2013. Viral Warfare: The Security Implications of Cyber and Biological Weapons. *Comparative Strategy*, 32(5):418–434. Publisher: Routledge _eprint: https://doi.org/10.1080/01495933.2013.821845. URL: https://doi.org/10.1080/01495933.2013.821845.

(53) Scott P. Layne, Tony J. Beugelsdijk, C. Kumar N. Patel, and National Academy of Sciences, editors. 2001. *Firepower in the lab: automation in the fight against infectious diseases and bioterrorism*. Henry, Washington, DC.

(54) National Research Council. 2011. *Challenges and Opportunities for Education About Dual Use Issues in the Life Sciences*. National Academies Press, Washington, D.C. URL: http://www.nap.edu/catalog/12958.

(55) Tristan A. Volpe. 2024. Biotechnology and the Dead Zone for Managing Dual-Use Dilemmas. Technical report, Nuclear Threat Initiative, Washington DC. URL: https://www.nti.org/analysis/articles/disincentivizing-bioweapons-theory-and-policy-approaches/#chapter8.

(56) David P. Clark and Nanette J. Pazdernik. 2016. Biological Warfare: Infectious Disease and Bioterrorism. *Biotechnology*, pages 687–719. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7150198/.

(57) David Hafemeister. 2016. Biological and Chemical Weapons. In *Nuclear Proliferation and Terrorism in the Post-9/11 World*, pages 337–351. Springer International Publishing. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7123302/.

(58) Brett Edwards, Tatyana Novossiolova, Michael Crowley, Simon Whitby, Malcolm Dando, and Lijun Shang. 2022. Meeting the Challenges of Chemical and Biological Weapons: Strengthening the Chemical and Biological Disarmament and Non-proliferation Regimes. *Frontiers in Political Science*, 4:805426. Publisher: Frontiers. URL: https://www.frontiersin.org/journals/political-science/articles/10.3389/fpos.2022.805426/full.

(59) 1986. Second review conference of the parties to the convention on the prohibition of the development, production and stockpiling of bacteriological (biological) and toxin weapons and on their destruction. In *United Nations Disarmament Yearbook 1986*, pages 263–287. UN. URL: https://www.un-ilibrary.org/disarmament/united-nations-disarmament-yearbook-1986_b6bb348f-en.

(60) Joseph M. Siracusa and Aiden Warren. 2018. The Nuclear Non-Proliferation Regime: An Historical Perspective. *Diplomacy & Statecraft*, 29(1):3–28. Publisher: Routledge _eprint: https://doi.org/10.1080/09592296.2017.1420495. URL: https://doi.org/10.1080/09592296.2017.1420495.

(61) Elisabeth Roehrlich. 2022. *Inspectors for Peace*. Johns Hopkins University Press. URL: https://www.press.jhu.edu/books/title/12352/inspectors-peace.

(62) Karen L. Mumy, William R. Howard, Ariel Parker, Jonathan Forman, and Gwyn Winfield. 2019. Organization for the Prohibition of Chemical Weapons (OPCW): History, Mission, and Accomplishments. In *Chemical Warfare Agents*, 3 edition. CRC Press. Num Pages: 11.

(63) William H. Tobey. 2018. A History of United Nations Security Council Resolution 1540. In Daniel Salisbury, Ian J. Stewart, and Andrea Viski, editors, *Preventing the Proliferation of WMDs: Measuring the Success of UN Security Council Resolution 1540*, pages 13–32. Springer International Publishing, Cham. URL: https://doi.org/10.1007/978-3-319-72203-0_2.

(64) Tyler Green. 2019. The Mechanisms Behind the United Nations Secretary-General's Mechanism – an Examination of the History and Application of the Fact-Finding Mechanism to the Alleged Chemical Attacks in Syria. *Journal of Biosecurity, Biosafety, and Biodefense Law*, 10.

(65) Matthew P. Shearer, Christina Potter, Rachel A. Vahey, Nancy D. Connell, and Gigi Kwik Gronvall. 2022. BWC assurance: increasing certainty in BWC compliance. *The Nonproliferation Review*, 29(1-3):47–75. URL: https://www.tandfonline.com/doi/full/10.1080/10736700.2023.2178099.

(66) Hans Blix. 1992. Verification of nuclear nonproliferation: The lesson of Iraq. *The Washington Quarterly*, 15(4):57–65. URL: http://www.tandfonline.com/doi/abs/10.1080/01636609209550118.

(67) John Carlson. 2006. Experience and Challenges in Weapons of Mass Destruction Treaty Verification: A Comparative View. In Rudolf Avenhaus, Nicholas Kyriakopoulos, Michel Richard, and Gotthard Stein, editors, *Verifying Treaty Compliance*, pages 213–234. Springer-Verlag, Berlin/Heidelberg. URL: http://link.springer.com/10.1007/3-540-33854-3_10.

(68) R Thakur and G Evans. 2013. *Nuclear weapons: the state of play*. Centre for Nuclear Non-Proliferation and Disarmament, Crawford School of Public Policy, The Australian National University. URL: http://hdl.handle.net/1885/10028.

(69) Steve Fetter. 1998. Verifying Nuclear Disarmament. In *Nuclear Weapons*. Routledge. Num Pages: 30.

(70) Michael Crowley, Malcolm Dando, and Lijun Shang, editors. 2018. *Preventing chemical weapons: arms control and disarmament as the sciences converge*. Royal Society of Chemistry, Cambridge.

(71) Gabriella Venturini. 2012. Control and Verification of Multilateral Treaties on Disarmament and Non-Proliferation of Weapons of Mass Destruction. *UC Davis J. Int'l L. & Pol'y*, 17(345). URL: https://jilp.law.ucdavis.edu/archives/17/2/control-and-verification-multilateral-treaties-disarmament-and-non-proliferation.

(72) Nicholas A. Sims and Jez Littlewood. 2011. Ambitious Incrementalism: Enhancing BWC Implementation in the Absence of a Verification Protocol. *The Nonproliferation Review*, 18(3):499–511. URL: http://www.tandfonline.com/doi/abs/10.1080/10736700.2011.618619.

(73) Matthew P. Shearer, Christina M. Potter, Rachel A. Vahey, Nicholas Munves, and Gigi Kwik Gronvall. 2025. BWC confidence-building measures: Increasing BWC assurance through transparency and information sharing. *Politics and the Life Sciences*, 44(1):5–27. URL: https://www.cambridge.org/core/product/identifier/S0730938424000091/type/journal_article.

(74) Paul Bernstien, Justin Anderson, Diane DiEuliis, Gerald Epstein, and Amanda Moodie. 2020. Weapons of Mass Destruction, Strategic Deterrence, and Great Power Competition. In *Strategic Assessment 2020 Into a New Era of Great Power Competition*. National Defense University Press, Washington DC. URL: https://ndupress.ndu.edu/Media/News/News-Article-View/Article/2404537/8-weapons-of-mass-destruction-strategic-deterrence-and-great-power-competition/.

(75) Graham Allison. 2008. Nuclear Deterrence in the Age of Nuclear Terrorism. *MIT Technology Review*, 111(6):68–73. URL: https://www.technologyreview.com/2008/10/20/218111/nuclear-deterrence-in-the-age-of-nuclear-terrorism-2/.

(76) Committee on Enhancing U.S. Nuclear Forensics and Attribution Support Capabilities, Committee on International Security and Arms Control, Policy and Global Affairs, and National Academies of Sciences, Engineering, and Medicine. 2021. *Restoring and Improving Nuclear Forensics to Support*

Attribution and Deterrence: Public Summary. National Academies Press, Washington, D.C. Pages: 26167. URL: https://www.nap.edu/catalog/26167.

(77) Elsa Hedling and Hedvig Ördén. 2025. Disinformation, deterrence and the politics of attribution. *International Affairs*, 101(3):967–986. URL: https://academic.oup.com/ia/article/101/3/967/8100244.

(78) Nathan Paxton and Jaime Yassif. 2024. Disincentivizing Bioweapons: Theory and Policy Approaches. Technical report, Nuclear Threat Initiative, Washington DC. URL: https://www.nti.org/analysis/articles/disincentivizing-bioweapons-theory-and-policy-approaches/.

(79) Sara M. Pires, Eric G. Evers, Wilfrid Van Pelt, Tracy Ayers, Elaine Scallan, Frederick J. Angulo, Arie Havelaar, and Tine Hald. 2009. Attributing the Human Disease Burden of Foodborne Infections to Specific Sources. *Foodborne Pathogens and Disease*, 6(4):417–424. URL: http://www.liebertpub.com/doi/10.1089/fpd.2008.0208.

(80) Christopher A Bidwell and Kishan Bhatt. 2016. Use of Attribution and Forensic Science in Addressing Biological Weapon Threats: A Multi-Faceted Study. Technical report, Federation of American Scientists. URL: https://uploads.fas.org/2016/03/NPSreport_final.pdf.

(81) The White House. 2022. National Biodefense Strategy and Implementation Plan: For Countering Biological Threats, Enhancing Pandemic Preparedness, and Achieving Global Health Security. Technical report, The White House, Washington DC. URL: https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/10/National-Biodefense-Strategy-and-Implementation-Plan-Final.pdf.

(82) Committee on Science Needs for Microbial Forensics: Developing an Initial International Roadmap, Board on Life Sciences, Division on Earth and Life Studies, and National Research Council. 2014. *Science Needs for Microbial Forensics: Developing Initial International Research Priorities*. National Academies Press (US), Washington (DC). URL: http://www.ncbi.nlm.nih.gov/books/NBK234876/.

(83) Jonathan B. Tucker and Gregory D. Koblentz. 2009. The Four Faces of Microbial Forensics. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, 7(4):389–397. URL: http://www.liebertpub.com/doi/10.1089/bsp.2009.0043.

(84) Kristina Daugirdas and Gian Luca Burci. 2019. Financing the World Health Organization: What Lessons for Multilateralism? *International Organizations Law Review*, 16(2):299–338. URL: https://brill.com/view/journals/iolr/16/2/article-p299_299.xml.

(85) Obichukwu Iwunna, Jonathan Kennedy, and Andrew Harmer. 2023. Flexibly funding WHO? An analysis of its donors' voluntary contributions. *BMJ Global Health*, 8(4):e011232. URL: https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2022-011232.

(86) Jin-Hong Yoo. 2025. On the Controversies Surrounding the Lab-Leak Theory of COVID-19. *Journal of Korean Medical Science*, 40(16):e153. URL: https://jkms.org/DOIx.php?id=10.3346/jkms.2025.40.e153.

(87) Kristian G. Andersen, Andrew Rambaut, W. Ian Lipkin, Edward C. Holmes, and Robert F. Garry. 2020. The proximal origin of SARS-CoV-2. *Nature Medicine*, 26(4):450–452. URL: http://www.nature.com/articles/s41591-020-0820-9.

(88) Peng Zhou, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, Yan Zhu, Bei Li, Chao-Lin Huang, Hui-Dong Chen, Jing Chen, Yun Luo, Hua Guo, Ren-Di Jiang, Mei-Qin Liu, Ying Chen, Xu-Rui Shen, Xi Wang, Xiao-Shuang Zheng, Kai Zhao, Quan-Jiao Chen, Fei Deng, Lin-Lin Liu, Bing Yan, Fa-Xian Zhan, Yan-Yi Wang, Geng-Fu Xiao, and Zheng-Li Shi. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798):270–273. URL: https://www.nature.com/articles/s41586-020-2012-7.

(89) Peter V. Markov, Mahan Ghafari, Martin Beer, Katrina Lythgoe, Peter Simmonds, Nikolaos I. Stilianakis, and Aris Katzourakis. 2023. The evolution of SARS-CoV-2. *Nature Reviews Microbiology*, 21(6):361–379. URL: https://www.nature.com/articles/s41579-023-00878-2.

(90) Edward C. Holmes. 2024. The Emergence and Evolution of SARS-CoV-2. *Annual Review of Virology*, 11(1):21–42. URL: https://www.annualreviews.org/content/journals/10.1146/annurev-virology-093022-013037.

(91) Isabel Pagani, Silvia Ghezzi, Simone Alberti, Guido Poli, and Elisa Vicenzi. 2023. Origin and evolution of SARS-CoV-2. *The European Physical Journal Plus*, 138(2):157. URL: https://link.springer.com/10.1140/epjp/s13360-023-03719-6.

(92) Alexander Crits-Christoph, Joshua I. Levy, Jonathan E. Pekar, Stephen A. Goldstein, Reema Singh, Zach Hensel, Karthik Gangavarapu, Matthew B. Rogers, Niema Moshiri, Robert F. Garry, Edward C. Holmes, Marion P. G. Koopmans, Philippe Lemey, Thomas P. Peacock, Saskia Popescu, Andrew Rambaut, David L. Robertson, Marc A. Suchard, Joel O. Wertheim, Angela L. Rasmussen, Kristian G. Andersen, Michael Worobey, and Florence Débarre. 2024. Genetic tracing of market wildlife and viruses at the epicenter of the COVID-19 pandemic. *Cell*, 187(19):5468–5482.e11. Publisher: Elsevier. URL: https://www.cell.com/cell/abstract/S0092-8674(24)00901-2.

(93) Joint WHO-China Study. 2021. WHO-convened global study of origins of SARS-CoV-2. Technical report, World Health Organization, Geneva, Switzerland. URL: https://www.who.int/publications-detail-redirect/who-convened-global-study-of-origins-of-sars-cov-2-china-part.

(94) W. Kip Viscusi. 2023. The global COVID-19 mortality cost report card: 2020, 2021, and 2022. *PLOS ONE*, 18(5):e0284273. URL: https://dx.plos.org/10.1371/journal.pone.0284273.

(95) Jesse D Bloom. 2021. Recovery of Deleted Deep Sequencing Data Sheds More Light on the Early Wuhan SARS-CoV-2 Epidemic. *Molecular Biology and Evolution*, 38(12):5211–5224. URL: https://academic.oup.com/mbe/article/38/12/5211/6353034.

(96) David A. Relman. 2020. To stop the next pandemic, we need to unravel the origins of COVID-19. *Proceedings of the National Academy of Sciences*, 117(47):29246–29248. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.2021133117.

(97) Rebecca Sohn. 2021. A Very Calm Guide to the Lab Leak Theory. *Slate*. Section: Medical Examiner. URL: https://slate.com/technology/2021/06/lab-leak-theory-questions-explainer.html.

(98) Alexandra Stevenson. 2020. Senator Tom Cotton Repeats Fringe Theory of Coronavirus Origins. *The New York Times*. URL: https://www.nytimes.com/2020/02/17/business/media/coronavirus-tom-cotton-china.html.

(99) Rund Abdelfatah. 2024. The story of a scientist who tried to stand for the truth and avoid Covid politics. *NPR*. URL: https://www.npr.org/2024/09/24/nx-s1-5121198/the-story-of-a-scientist-who-tried-to-stand-for-the-truth-and-avoid-covid-politics.

(100) L.P. Villarreal. 2008. Evolution of Viruses. In *Encyclopedia of Virology*, pages 174–184. Elsevier. URL: https://linkinghub.elsevier.com/retrieve/pii/B9780123744104007068.

(101) Siobain Duffy, Laura A. Shackelton, and Edward C. Holmes. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics*, 9(4):267–276. URL: https://www.nature.com/articles/nrg2323.

(102) Bryan T. Grenfell, Oliver G. Pybus, Julia R. Gog, James L. N. Wood, Janet M. Daly, Jenny A. Mumford, and Edward C. Holmes. 2004. Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. *Science*, 303(5656):327–332. URL: https://www.science.org/doi/10.1126/science.1090727.

(103) Esteban Domingo, Julie Sheldon, and Celia Perales. 2012. Viral Quasispecies Evolution. *Microbiology and Molecular Biology Reviews*, 76(2):159–216. URL: https://journals.asm.org/doi/10.1128/MMBR.05023-11.

(104) Jocelyn Kaiser. 2024. 'Lab-leak' proponents at Rutgers accused of defaming and intimidating COVID-19 origin researchers. *Science Insider*. URL: https://www.science.org/

content/article/lab-leak-proponents-rutgers-accused-defaming-and-intimidating-covid-19-origin.

(105) Susan Maret. 2025. Introduction, Special Issue on Pandemic Secrecy: The COVID Origin Story and Pandemic Risk Society. *Secrecy and Society*, 3(2). URL: https://scholarworks.sjsu.edu/secrecyandsociety/vol3/iss2/1.

(106) Select Subcommittee on the Coronavirus Pandemic. 2025. After Action Review of the COVID-19 Pandemic: The Lessons Learned and a Path Forward. Technical report, Washington DC. URL: https://oversight.house.gov/report/after-action-review-of-the-covid-19-pandemic-the-lessons-learned-and-a-path-forward/.

(107) Office of the Director National Intelligence. 2025. Unclassified Summary of Assessment on COVID-19 Origins. Technical report, Washington DC. URL: https://www.dni.gov/files/ODNI/documents/assessments/Unclassified-Summary-of-Assessment-on-COVID-19-Origins.pdf.

(108) Gregory D. Koblentz and Rocco Casagrande. 2024. Beyond gain of function: strengthening oversight of research with potential pandemic pathogens. *Pathogens and Global Health*, 118(3):197–208.

(109) Kelsey Lane Warmbrod, Michael G Montague, and Gigi Kwik Gronvall. 2021. COVID-19 and the gain of function debates: Improving biosafety measures requires a more precise definition of which experiments would raise safety concerns. *EMBO reports*, 22(10):e53739. URL: https://www.embopress.org/doi/10.15252/embr.202153739.

(110) John P. Moore. 2021. Is the debate over the origin of Covid-19 still worth having? URL: https://www.statnews.com/2021/11/05/is-debate-about-origin-of-covid-19-still-worth-having/.

(111) 2025. WATCH LIVE: Trump addresses UN General Assembly for first time since reelection. URL: https://www.youtube.com/watch?v=9iMAKayuz4U.

(112) Julia Taliesin. 2021. Harvard professor signs international letter calling for investigation into COVID-

19 origins. URL: https://www.boston.com/news/coronavirus/2021/05/17/harvard-professor-investigation-coronavirus-origins-marc-lipsitch/.

(113) Tom Inglesby. 2025. Tom Inglesby posted on the topic COVID Origins and Effects: A New Report by Bob Kadlec. URL: https://www.linkedin.com/posts/tom-inglesby-93b51582_over-12-million-americans-died-by-covid-activity-7355629844002471937-zaVD.

(114) Marc Lipsitch and Thomas V. Inglesby. 2014. Moratorium on research intended to create novel potential pandemic pathogens. *mBio*, 5(6):e02366–14.

(115) The White House. 2025. Lab Leak: The True Origins of Covid-19. Technical report, The White House, Washington DC. URL: https://www.whitehouse.gov/lab-leak-true-origins-of-covid-19/.

(116) Jonathan E. Pekar, Andrew Magee, Edyth Parker, Niema Moshiri, Katherine Izhikevich, Jennifer L. Havens, Karthik Gangavarapu, Lorena Mariana Malpica Serrano, Alexander Crits-Christoph, Nathaniel L. Matteson, Mark Zeller, Joshua I. Levy, Jade C. Wang, Scott Hughes, Jungmin Lee, Heedo Park, Man-Seong Park, Katherine Ching Zi Yan, Raymond Tzer Pin Lin, Mohd Noor Mat Isa, Yusuf Muhammad Noor, Tetyana I. Vasylyeva, Robert F. Garry, Edward C. Holmes, Andrew Rambaut, Marc A. Suchard, Kristian G. Andersen, Michael Worobey, and Joel O. Wertheim. 2022. The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2. *Science*, 377(6609):960–966. URL: https://www.science.org/doi/10.1126/science.abp8337.

(117) P. M. Sharp and B. H. Hahn. 2011. Origins of HIV and the AIDS Pandemic. *Cold Spring Harbor Perspectives in Medicine*, 1(1):a006841–a006841. URL: http://perspectivesinmedicine.cshlp.org/lookup/doi/10.1101/cshperspect.a006841.

(118) Jie Cui, Fang Li, and Zheng-Li Shi. 2019. Origin and evolution of pathogenic coronaviruses. *Nature Reviews Microbiology*, 17(3):181–192. URL: https://www.nature.com/articles/s41579-018-0118-9.

(119) Robert Belshe. 2005. The Origins of Pandemic Influenza — Lessons from the 1918 Virus. *New England Journal of Medicine*, 353(21):2209–2211. URL: https://www.nejm.org/doi/full/10.1056/NEJMp058281.

(120) Rebecca J. Garten, C. Todd Davis, Colin A. Russell, Bo Shu, Stephen Lindstrom, Amanda Balish, Wendy M. Sessions, Xiyan Xu, Eugene Skepner, Varough Deyde, Margaret Okomo-Adhiambo, Larisa Gubareva, John Barnes, Catherine B. Smith, Shannon L. Emery, Michael J. Hillman, Pierre Rivailler, James Smagala, Miranda De Graaf, David F. Burke, Ron A. M. Fouchier, Claudia Pappas, Celia M. Alpuche-Aranda, Hugo López-Gatell, Hiram Olivera, Irma López, Christopher A. Myers, Dennis Faix, Patrick J. Blair, Cindy Yu, Kimberly M. Keene, P. David Dotson, David Boxrud, Anthony R. Sambol, Syed H. Abid, Kirsten St. George, Tammy Bannerman, Amanda L. Moore, David J. Stringer, Patricia Blevins, Gail J. Demmler-Harrison, Michele Ginsberg, Paula Kriner, Steve Waterman, Sandra Smole, Hugo F. Guevara, Edward A. Belongia, Patricia A. Clark, Sara T. Beatrice, Ruben Donis, Jacqueline Katz, Lyn Finelli, Carolyn B. Bridges, Michael Shaw, Daniel B. Jernigan, Timothy M. Uyeki, Derek J. Smith, Alexander I. Klimov, and Nancy J. Cox. 2009. Antigenic and Genetic Characteristics of Swine-Origin 2009 A(H1N1) Influenza Viruses Circulating in Humans. *Science*, 325(5937):197–201. URL: https://www.science.org/doi/10.1126/science.1176225.

(121) Mark P. Zwart and Santiago F. Elena. 2015. Matters of Size: Genetic Bottlenecks in Virus Infection and Their Potential Impact on Evolution. *Annual Review of Virology*, 2(1):161–179. URL: https://www.annualreviews.org/doi/10.1146/annurev-virology-100114-055135.

(122) John T McCrone and Adam S Lauring. 2018. Genetic bottlenecks in intraspecies virus transmission. *Current Opinion in Virology*, 28:20–25. URL: https://linkinghub.elsevier.com/retrieve/pii/S1879625717300779.

(123) Lijuan Yin, Shuli Man, Shengying Ye, Guozhen Liu, and Long Ma. 2021. CRISPR-Cas based virus detection: Recent advances and perspectives. *Biosensors and Bioelectronics*, 193:113541. URL: https://linkinghub.elsevier.com/retrieve/pii/S0956566321005789.

(124) Andreas S. Puschnik, Karim Majzoub, Yaw Shin Ooi, and Jan E. Carette. 2017. A CRISPR toolbox

to study virus–host interactions. *Nature Reviews Microbiology*, 15(6):351–364. URL: https://www.nature.com/articles/nrmicro.2017.29.

(125) Catherine A. Freije and Pardis C. Sabeti. 2021. Detect and destroy: CRISPR-based technologies for the response against viruses. *Cell Host & Microbe*, 29(5):689–703. URL: https://linkinghub.elsevier.com/retrieve/pii/S1931312821001517.

(126) Jasper Adriaan Soppe and Robert Jan Lebbink. 2017. Antiviral Goes Viral: Harnessing CRISPR/-Cas9 to Combat Viruses in Humans. *Trends in Microbiology*, 25(10):833–850. URL: https://linkinghub.elsevier.com/retrieve/pii/S0966842X17300938.

(127) Alexander T. Ciota, Amy O. Lovelace, Kiet A. Ngo, An N. Le, Joseph G. Maffei, Mary A. Franke, Anne F. Payne, Susan A. Jones, Elizabeth B. Kauffman, and Laura D. Kramer. 2007. Cell-specific adaptation of two flaviviruses following serial passage in mosquito cell culture. *Virology*, 357(2):165–174. URL: https://linkinghub.elsevier.com/retrieve/pii/S0042682206005496.

(128) Marianna Sockrider, Shazia Jamil, Lekshmi Santhosh, and W. Graham Carlos. 2020. COVID-19 Infection versus Influenza (Flu) and Other Respiratory Illnesses. *American Journal of Respiratory and Critical Care Medicine*, 202(10):P27–P28. URL: https://www.atsjournals.org/doi/10.1164/rccm.2020C16.

(129) World Health Organization. (2025). Independent assessment of the origins of SARS-CoV-2. 2025. Independent assessment of the origins of SARS-CoV-2, developed by the Scientific Advisory Group for the Origins of Novel Pathogens (SAGO). Technical report, World Health Organization, Geneva, Switzerland. URL: https://www.who.int/publications/m/item/independent-assessment-of-the-origins-of-sars-cov-2-from-the-scientific-advisory-group-for-the-origins-of-novel-pathogens.

(130) Scientific Advisory Group for the Origins of Novel Pathogens (SAGO). 2022. Preliminary Report of the SAGO. Technical report, World Health Organization, Geneva, Switzerland. URL:

https://cdn.who.int/media/docs/default-source/scientific-advisory-group-on-the-origins-of-novel-pathogens/sago-report-09062022.pdf.

(131) Anne E. Sundelson, Amelia M. Jamison, Noelle Huhn, Sarah-Louise Pasquino, and Tara Kirk Sell. 2023. Fighting the infodemic: the 4 i Framework for Advancing Communication and Trust. *BMC Public Health*, 23(1):1662. Publisher: BioMed Central. URL: https://link.springer.com/article/10.1186/s12889-023-16612-9.

(132) Tara Kirk Sell, Divya Hosangadi, and Marc Trotochaud. 2020. Misinformation and the US Ebola communication crisis: analyzing the veracity and content of social media messages related to a fear-inducing infectious disease outbreak. *BMC Public Health*, 20(1):550. Publisher: BioMed Central. URL: https://link.springer.com/article/10.1186/s12889-020-08697-3.

(133) Tara Kirk Sell, Sanjana J. Ravi, Crystal Watson, Diane Meyer, Laura E. Pechta, Dale A. Rose, Keri M. Lubell, Michelle N. Podgornik, and Monica Schoch-Spana. 2020. A Public Health Systems View of Risk Communication About Zika. *Public Health Reports®*, 135(3):343–353. URL: https://journals.sagepub.com/doi/10.1177/0033354920912215.

(134) Boghuma K. Titanji, Isabelle Mekone, David Scales, Shalom Tchokfe Ndoula, Judith Seungue, and Sara Gorman. Pre-emptively tackling vaccine misinformation for a successful large-scale roll-out of malaria vaccines in Africa. URL: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(23)00453-X/abstract.

(135) Siv Hilde Berg, Jane K. O'Hara, Marie Therese Shortt, Henriette Thune, Kolbjørn Kallesten Brønnick, Daniel Adrian Lungu, Jo Røislien, and Siri Wiig. 2021. Health authorities' health risk communication with the public during pandemics: a rapid scoping review. *BMC Public Health*, 21(1):1401. URL: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021-11468-3.

(136) Nashit Chowdhury, Ayisha Khalid, and Tanvir C. Turin. 2021. Understanding misinformation infodemic during public health emergencies due to large-scale disease outbreaks: a rapid review. *Journal*

of Public Health, 31(4):553–573. Publisher: Springer. URL: https://link.springer.com/article/10.1007/s10389-021-01565-3.

(137) Ho-Chun Herbert Chang and Emilio Ferrara. 2022. Comparative analysis of social bots and humans during the COVID-19 pandemic. *Journal of Computational Social Science*, 5(2):1409–1425. Publisher: Springer. URL: https://link.springer.com/article/10.1007/s42001-022-00173-9.

(138) Gourja Bansal, Kiran Narta, and Manoj Ramesh Teltumbade. 2018. Next-Generation Sequencing: Technology, Advancements, and Applications. In Asheesh Shanker, editor, *Bioinformatics: Sequences, Structures, Phylogeny*, pages 15–46. Springer, Singapore. URL: https://doi.org/10.1007/978-981-13-1562-6_2.

(139) Mei-Ling Han, Yan Zhu, Darren J. Creek, Yu-Wei Lin, Alina D. Gutu, Paul Hertzog, Tony Purcell, Hsin-Hui Shen, Samuel M. Moskowitz, Tony Velkov, and Jian Li. 2019. Comparative Metabolomics and Transcriptomics Reveal Multiple Pathways Associated with Polymyxin Killing in Pseudomonas aeruginosa. *mSystems*, 4(1):10.1128/msystems.00149–18. URL: https://journals.asm.org/doi/10.1128/msystems.00149-18.

(140) Arizaldo E. Castro and Maria Corazon A. De Ungria. 2022. Methods used in Microbial Forensics and Epidemiological Investigations for Stronger Health Systems. *Forensic Sciences Research*, 7(4):650–661. URL: https://academic.oup.com/fsr/article/7/4/650-661/7071966.

(141) Manuela Oliveira, Kamila Marszałek, Michał Kowalski, Alina Frolova, Paweł P. Łabaj, Wojciech Branicki, Áurea Madureira-Carvalho, Diana Dias Da Silva, and Ricardo Jorge Dinis-Oliveira. 2024. Sequencing Technologies in Forensic Microbiology: Current Trends and Advancements. *Forensic Sciences*, 4(4):523–545. URL: https://www.mdpi.com/2673-6756/4/4/35.

(142) Irene Kuiper. 2016. Microbial forensics: next-generation sequencing as catalyst: The use of new sequencing technologies to analyze whole microbial communities could become a powerful tool for forensic and criminal investigations. *EMBO reports*, 17(8):1085–1087.

(143) Ana Sofia Ribeiro Duarte, Timo Röder, Liese Van Gompel, Thomas Nordahl Petersen, Rasmus Borup Hansen, Inge Marianne Hansen, Alex Bossers, Frank M. Aarestrup, Jaap A. Wagenaar, and Tine Hald. 2021. Metagenomics-Based Approach to Source-Attribution of Antimicrobial Resistance Determinants – Identification of Reservoir Resistome Signatures. *Frontiers in Microbiology*, 11:601407. URL: https://www.frontiersin.org/articles/10.3389/fmicb.2020.601407/full.

(144) J.s. West and R.b.e. Kimber. 2015. Innovations in air sampling to detect plant pathogens. *Annals of Applied Biology*, 166(1):4–17. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/aab.12191. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/aab.12191.

(145) David Bass, Kevin W. Christison, Grant D. Stentiford, Lauren S. J. Cook, and Hanna Hartikainen. 2023. Environmental DNA/RNA for pathogen and parasite detection, surveillance, and ecology. *Trends in Parasitology*, 39(4):285–304. Publisher: Elsevier. URL: https://www.cell.com/trends/parasitology/abstract/S1471-4922(22)00315-4.

(146) Shirin Alex, Thomas P. Shehata, Andreea Iris Gergely, and Marcel de Puit. 2025. Proteomics in forensics: from source attribution to reconstruction of events. *Science & Justice*, 65(6):101320. URL: https://www.sciencedirect.com/science/article/pii/S1355030625001042.

(147) Michal Szeremeta, Karolina Pietrowska, Anna Niemcunowicz-Janica, Adam Kretowski, and Michal Ciborowski. 2021. Applications of Metabolomics in Forensic Toxicology and Forensic Medicine. *International Journal of Molecular Sciences*, 22(6):3010. Publisher: Multidisciplinary Digital Publishing Institute. URL: https://www.mdpi.com/1422-0067/22/6/3010.

(148) Atif Khurshid Wani, Nahid Akhtar, Tahir ul Gani Mir, Chirag Chopra, Reena Singh, Jong Chan Hong, and Ulhas Sopanrao Kadam. 2024. CRISPR/Cas12a-based biosensors for environmental monitoring and diagnostics. *Environmental Technology & Innovation*, 34:103625. URL: https://www.sciencedirect.com/science/article/pii/S2352186424001019.

(149) Gertjan Medema, Leo Heijnen, Goffe Elsinga, Ronald Italiaander, and Anke Brouwer. 2020. Presence

of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in The Netherlands. *Environmental Science & Technology Letters*, 7(7):511–516. Publisher: American Chemical Society. URL: https://doi.org/10.1021/acs.estlett.0c00357.

(150) Lindsay Morton, Kathleen Creppage, Nazia Rahman, June Early, Laurie Hartman, Ashley Hydrick, and Matthew Kasper. 2024. Challenges and Opportunities in Pathogen Agnostic Sequencing for Public Health Surveillance: Lessons Learned From the Global Emerging Infections Surveillance Program. *Health Security*, 22(1):16–24.

(151) Karrie K. K. Ko, Kern Rei Chng, and Niranjan Nagarajan. 2022. Metagenomics-enabled microbial surveillance. *Nature Microbiology*, 7(4):486–496. URL: https://www.nature.com/articles/s41564-022-01089-w.

(152) Amesh A. Adalja, Eric Toner, and Thomas V. Inglesby. 2020. Priorities for the US Health Community Responding to COVID-19. *JAMA*, 323(14):1343. URL: https://jamanetwork.com/journals/jama/fullarticle/2762690.

(153) Dana A Shea and Sarah A Lister. 2003. The BioWatch Program: Detection of Bioterrorism. Technical Report Report No. RL 32152, Congressional Research Service, Washington DC.

(154) Aaron Adler, Joel S. Bader, Brian Basnight, Benjamin W. Booth, Jitong Cai, Elizabeth Cho, Joseph H. Collins, Yuchen Ge, John Grothendieck, Kevin Keating, Tyler Marshall, Anton Persikov, Helen Scott, Roy Siegelmann, Mona Singh, Allison Taggart, Benjamin Toll, Kenneth H. Wan, Daniel Wyschogrod, Fusun Yaman, Eric M. Young, Susan E. Celniker, and Nicholas Roehner. 2024. Ensemble Detection of DNA Engineering Signatures. *ACS Synthetic Biology*, 13(4):1105–1115. URL: https://pubs.acs.org/doi/10.1021/acssynbio.3c00398.

(155) Vivienne Machi. 2018. IARPA-Funded Initiatives Could Thwart Biothreats. *National Defense Magazine*. URL: https://www.nationaldefensemagazine.org/articles/2018/8/1/iarpa-funded-initiatives-could-thwart-biothreats.

(156) Gene Godbold, Jody Proescher, and Pascale Gaudet. 2025. New and revised gene ontology biological process terms describe multiorganism interactions critical for understanding microbial pathogenesis and sequences of concern. *Journal of Biomedical Semantics*, 16(1):4. URL: https://jbiomedsem.biomedcentral.com/articles/10.1186/s13326-025-00323-8.

(157) Gene D. Godbold, F. Curtis Hewitt, Anthony D. Kappell, Matthew B. Scholz, Stacy L. Agar, Todd J. Treangen, Krista L. Ternus, Jonas B. Sandbrink, and Gregory D. Koblentz. 2023. Improved understanding of biorisk for research involving microbial modification using annotated sequences of concern. *Frontiers in Bioengineering and Biotechnology*, 11:1124100. URL: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1124100/full.

(158) Deqiao Tian and Tao Zheng. 2014. Comparison and Analysis of Biological Agent Category Lists Based On Biosafety and Biodefense. *PLoS ONE*, 9(6):e101163. URL: https://dx.plos.org/10.1371/journal.pone.0101163.

(159) 2017. Working Paper submitted by Canada, the United Kingdom of Great Britain and Northern Ireland and the United States of America, to the Meeting of the States Parties to the Convention on the Prohibition of the Development, Production and Stockpiling of Bacteriological (Biological) and Toxin Weapons and on Their Destruction. Technical Report WC/MSP/2017/WP.20, Biological and Toxins Weapons Convention, Geneva, Switzerland.

(160) UN General Assembly. 1987. Resolution 42/37C. Technical Report A/RES/42/37C, United Nations, New York City.

(161) UN General Assembly. 2006. Resolution 60/288. Technical Report A/RES/60/288, United Nations, New York City.

(162) UN Office for Disarmament Affairs. Secretary-General's Mechanism for Investigation of Alleged Use of Chemical and Biological Weapons (UNSGM) | United Nations Office for Disarmament Affairs. URL: https://disarmament.unoda.org/en/our-work/weapons-mass-destruction/secretary-generals-mechanism-investigation-alleged-use-chemical.

(163) Michael Worobey, Joshua I. Levy, Lorena Malpica Serrano, Alexander Crits-Christoph, Jonathan E. Pekar, Stephen A. Goldstein, Angela L. Rasmussen, Moritz U. G. Kraemer, Chris Newman, Marion P. G. Koopmans, Marc A. Suchard, Joel O. Wertheim, Philippe Lemey, David L. Robertson, Robert F. Garry, Edward C. Holmes, Andrew Rambaut, and Kristian G. Andersen. 2022. The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. *Science*, 377(6609):951–959. Publisher: American Association for the Advancement of Science. URL: https://www.science.org/doi/10.1126/science.abp8715.

(164) Steven Lee Myers and Rick Gladstone. 2013. Russia Calls U.N. Chemical Report on Syria Biased. *The New York Times*. URL: https://www.nytimes.com/2013/09/19/world/middleeast/syria.html.

(165) James D. Haslam, Paul Russell, Stephanie Hill, Stevan R. Emmett, and Peter G. Blain. 2022. Chemical, biological, radiological, and nuclear mass casualty medicine: a review of lessons from the Salisbury and Amesbury Novichok nerve agent incidents. *British Journal of Anaesthesia*, 128(2):e200–e205.

(166) Elisabetta Mereu, Atefeh Lafzi, Catia Moutinho, Christoph Ziegenhain, Davis J. McCarthy, Adrián Álvarez Varela, Eduard Batlle, Sagar, Dominic Grün, Julia K. Lau, Stéphane C. Boutet, Chad Sanada, Aik Ooi, Robert C. Jones, Kelly Kaihara, Chris Brampton, Yasha Talaga, Yohei Sasagawa, Kaori Tanaka, Tetsutaro Hayashi, Caroline Braeuning, Cornelius Fischer, Sascha Sauer, Timo Trefzer, Christian Conrad, Xian Adiconis, Lan T. Nguyen, Aviv Regev, Joshua Z. Levin, Swati Parekh, Aleksandar Janjic, Lucas E. Wange, Johannes W. Bagnoli, Wolfgang Enard, Marta Gut, Rickard Sandberg, Itoshi Nikaido, Ivo Gut, Oliver Stegle, and Holger Heyn. 2020. Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nature Biotechnology*, 38(6):747–755. Publisher: Nature Publishing Group. URL: https://www.nature.com/articles/s41587-020-0469-4.

(167) Silvie Van den Hoecke, Judith Verhelst, Marnik Vuylsteke, and Xavier Saelens. 2015. Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing. *BMC Genomics*, 16(1):79. URL: https://doi.org/10.1186/s12864-015-1284-z.

(168) Fatos Selita, Vanessa Smereczynska, Robert Chapman, Teemu Toivainen, and Yulia Kovas. 2020.

Judging in the genomic era: judges' genetic knowledge, confidence and need for training. *European Journal of Human Genetics*, 28(10):1322–1330. Publisher: Nature Publishing Group. URL: https://www.nature.com/articles/s41431-020-0650-8.

(169) Eric D. Merkley. 2019. Proteomics for Microbial Forensics. In *Applications in Forensic Proteomics: Protein Identification and Profiling*, volume 1339 of *ACS Symposium Series*, pages 143–160. American Chemical Society. Section: 9. URL: https://doi.org/10.1021/bk-2019-1339.ch009.

(170) David Wunschel, Edan Tulman, Heather Engelmann, Brian H. Clowers, Steven Geary, Aaron Robinson, and Xiaofen Liao. 2013. Forensic proteomics of poxvirus production. *The Analyst*, 138(21):6385–6397.

(171) M. E. Kosal and D. E. Anderson. 2004. An unaddressed issue of agricultural terrorism: A case study on feed security1. *Journal of Animal Science*, 82(11):3394–3400. URL: https://doi.org/10.2527/2004.82113394x.

(172) National Institute of Justice. 2023. Crime Scene and DNA Basics for Forensic Analysts | Types of Evidence. URL: https://nij.ojp.gov/nij-hosted-online-training-courses/crime-scene-and-dna-basics-forensic-analysts/evidence-crime-scene/types-evidence.

(173) Kinjiro Morimoto, D. Craig Hooper, Heather Carbaugh, Zhen Fang Fu, Hilary Koprowski, and Bernhard Dietzschold. 1998. Rabies virus quasispecies: Implications forpathogenesis. *Proceedings of the National Academy of Sciences*, 95(6):3152–3156. Publisher: Proceedings of the National Academy of Sciences. URL: https://www.pnas.org/doi/10.1073/pnas.95.6.3152.

(174) John Cohen. 2023. U.S. cancels program aimed at identifying potential pandemic viruses. URL: https://www.science.org/content/article/u-s-cancels-program-aimed-identifying-potential-pandemic-viruses.

(175) Alexandria B Boehm, Bridgette Hughes, Dorothea Duong, Vikram Chan-Herur, Anna Buchman, Marlene K Wolfe, and Bradley J White. 2023. Wastewater concentrations of human in-

fluenza, metapneumovirus, parainfluenza, respiratory syncytial virus, rhinovirus, and seasonal coronavirus nucleic-acids during the COVID-19 pandemic: a surveillance study. *The Lancet Microbe*, 4(5):e340–e348. URL: https://www.sciencedirect.com/science/article/pii/S266652472200386X.

(176) Jocelyn Kaiser. 2024. International panel calls for tighter oversight of risky pathogen studies. URL: https://www.science.org/content/article/international-panel-calls-tighter-oversight-risky-pathogen-studies.

(177) Fei Wang, Pan Li, Hoi Ching Chu, and Pik Kwan Lo. 2022. Nucleic Acids and Their Analogues for Biomedical Applications. *Biosensors*, 12(2):93. Publisher: Multidisciplinary Digital Publishing Institute. URL: https://www.mdpi.com/2079-6374/12/2/93.

(178) Paul C. Stern. 2000. Psychology and the science of human-environment interactions. *American Psychologist*, 55(5):523–530. Place: US Publisher: American Psychological Association.

(179) Leonard Krasner. 1980. *Environmental design and human behavior: a psychology of the individual in society*. Number v. 85 in Pergamon general psychology series. Pergamon Press, New York.

(180) Charles S. Carver and Michael F. Scheier. 1982. Control theory: A useful conceptual framework for personality–social, clinical, and health psychology. *Psychological Bulletin*, 92(1):111–135. Place: US Publisher: American Psychological Association.

(181) Steven Taylor. 2022. The psychology of pandemics: Lessons learned for the future. *Canadian Psychology / Psychologie canadienne*, 63(2):233–246. Place: US Publisher: Educational Publishing Foundation.

(182) American Public Health Association. 2025. Public Health Under Threat. URL: https://www.apha.org/topics-and-issues/public-health-under-threat.

(183) Brandy Zadrozny. 2025. CDC is pulling back $11B in Covid funding sent to health departments across the U.S. URL: https://www.nbcnews.com/health/health-news/cdc-pulling-back-11b-covid-funding-sent-health-departments-us-rcna198006.

(184) Mike Stobbe and Lauren Neergaard. 2025. Kennedy's vaccine advisers change COVID shot guidance, calling them an individual choice. Section: Health. URL: https://apnews.com/article/cdc-advisers-vaccines-hepatitis-b-covid19-a1fb61ef7e29d93d2828618d890cceb3.

Chapter A

Appendix

A.1 Scenario 1

Death of James Salle

A.1.1 Text of Scenario

Background James Salle was the billionaire CEO of HSH, a multinational oil and diamond company headquartered in the country of Xenon. Salle was an international celebrity who made inflammatory comments on social media, had flashy relationships with models half his age, and was often accused of questionable business dealings. Recently, Salle was the subject of protests after announcing HSH plans to build an oil line through several countries. Construction of the line would displace Indigenous populations and, according to some environmental experts, place 3 bodies of water vital to local economies and livelihoods at high risk of pollution. Salle was rumored to give significant amounts of money to national-level politicians set to vote on whether to allow HSH to build the oil line. Politicians from adjacent countries strongly opposed the project and condemned HSH and Salle. Environmental groups with a history of violent attacks against other corporations, as well as several anonymous online figures, made threats against Salle and HSH.

April 5: James Salle was speaking at an outdoor environmental stewardship event in Argon Capital City, Argon, when a dart was shot from the audience area and hit Salle, piercing his upper right arm. The event was immediately stopped, but a perpetrator(s) was not found.

May 3: Exactly 28 days after the event, HSH announced James Salle died of rabies in Xenon, where he lived.

Theories of how Salle contracted rabies began to circulate. Several prominent figures publicly suggested Salle was assassinated with a bioweapon/rabies laced dart. Officials from Xenon publicly called on the event host nation, Argon, to share evidence on the dart incident.

May 11: Argon shared the collected evidence with Xenon. The information from the Argon Capital City Police Department showed the dart that hit Salle had genetic fragments of rabies virus on its surface. The evidence was quickly leaked, prompting international calls to investigate the potential use of a bioweapon for an assassination and identify the perpetrator(s).

May 12: Xenon and Argon announced they created a 16-member joint task force to investigate the event. Over the course of the investigation, the Joint Argon- Xenon investigatory team visited Argon Capitol City and collected surveillance video footage, environmental samples, and witness statements, as well as evidence previously gathered by the Argon Police Department. They also traveled to Xenon, where James Salle died. Working with the hospital and medical examiner, the team examined post-mortem analysis and previously collected samples.

Evidence

1. Event venue video surveillance. Video recording of the event shows the stage and Salle being hit by the dart but does not show who threw the dart. All entrances and exits were recorded throughout the event. Using various sources—surveillance footage, facial recognition software, social media profiles, visas, and national databases from Xenon and Argon—the team identified 369 of the 437 confirmed event attendees who entered or exited the venue on April 5. Overall, 68 event attendees were not identified on footage from the venue.

2. Surveillance footage from Argon Capital City CCTV. The movements of the 68 people who were in the camera blind spots immediately before and after the dart hit Salle were identified using Argon Capital City CCTV footage from April 5. Based on the video, none of those individuals engaged in activity deemed suspicious that day.

3. Environmental sampling of venue. The team swabbed the stage, entrances, exits, and handrails in the event venue. More than 200 samples underwent PCR testing to detect rabies nucleic acid, but none

of the tests were positive. Notably, this environmental sampling occurred 43 days after Salle died. This environmental testing was conducted to assess if there was an ongoing source of rabies virus at the venue.

4. Witness Statements. The Argon Capital City Police Department interviewed attendees— including venue staff and event organizers. Eight witnesses near where the dart was likely to have originated reported seeing a person in a dark hoodie running by them around the time Salle was hit but none could provide additional details on the person's identity.

5. Medical records from the hospital where Salle died. Salle's medical records show he was diagnosed with rabies and medical care consisted of palliative care. The records confirm Salle died of rabies. The records also contain the genetic sequencing results of the samples collected at the hospital.

6. Post-mortem samples collected by Xenon Medical Examiner. Cerebrospinal fluid (CSF), serum, and saliva samples from Salle's body were sequenced for rabies virus and tested positive. The medical examiner shared both the samples and the sequencing results with the investigatory team.

7. Post-mortem samples reanalysis. The investigatory team conducted its own analysis of the previously collected CSF, serum, and saliva samples from Salle's body. They sequenced the rabies virus and conducted metagenomic sequencing. The sequencing results from the investigatory team's sequencing were 98.7% similar to the results from the Xenon Medical Examiner's office sequencing. The sequences collected from these autopsy samples were used to generate a consensus sequence for comparative analysis to other samples collected later in the investigation.

8. Genomic analysis of post-mortem samples. The investigatory team's bioinformaticist and virologist analyzed all the sequences collected from samples taken from Salle. Analysis indicated there were 2 copies of the rabies virus glycoprotein (G) gene present in more than 90% of the virus genomes sequenced, suggesting the virus had been genetically engineered to contain 2 copies of the G gene rather than the single copy that is present in rabies virus samples isolated to date. Based on prior rabies virus research, 2 copies of the G protein may increase infectivity. Sequencing also indicated that over 70% of the sequences had mutations in both copies of the G gene at 4 sites that papers in the literature report lead to enhance viral pathogenesis.

9. Metagenomic sequencing of swabs from dart. The investigatory team conducted its own testing of the dart, confirming the Police Department's findings that found fragments of rabies nucleic acid. Those

fragments are sequenced, finding a 99.2% nucleotide sequence similarity with the consensus sequence created from the autopsy samples. Salle's DNA also is detected on the dart.

10. Interviews with Salle's personal staff and close contacts. The investigatory team interviewed members of Salle's personal and household staff, including his personal assistant and house manager, as well as his close contacts. None of these individuals reported Salle mentioning or exhibiting symptoms of illness prior to his speech. Additionally, none of the individuals said they were aware of Salle being bit by any animals in the months preceding his death.

11. Social media analysis. The investigatory team analyzed social media posts from the event, including videos and photos. No image was found of someone throwing a dart. Of the 437 attendees, 148 had made positive or negative public comments about James Salle or HSH. Additionally, 53 attendees had posted using broadly threatening language toward a person, government, or company. Of the 68 people in camera blind spots when Salle was attacked, 2 previously posted disparaging comments about James Salle or HSH: Jim Little and Roy Smith.

The investigatory team identified 2 suspects: Roy Smith and Jim Little. Details about these individuals and subsequent evidence from investigations are as follows.

Information collected about Roy Smith: Roy Smith is a communications director for a non-profit organization that aims to save animals from climate change threats. He studied molecular biology and communications as an undergraduate in college and previously worked as a technician in a tuberculosis lab. He lives in a large city in Xenon with 2 male roommates who work in laboratories. The facilities employing Smith's roommates, Biotech Facility and University Lab, allowed the investigatory team to conduct inspections. It is noted that Smith's cousin, who lived in the same area, died of rabies 2 years earlier after being bitten by a fox.

Information collected about Jim Little: Jim Little is a scientist in a pharmacogenetics laboratory that specializes in hepatitis A treatments at the Argon Institute of Medicine (IOM). Little has worked with live viruses and targeted viral gene modification. He reports being friends with Jasper Bill, a researcher who worked in the IOM's rabies lab but who died earlier this year in a car accident. The investigatory team inspected the IOM facility, including its security systems. The upper floors of the building, where all the labs are located, require the use of ID badges to enter and are closely monitored via security cameras.

Employees are only given access to floors with their own labs or shared equipment.

Specific evidence on Roy Smith:

12. Interview with Roy Smith. Smith was aggressive in his interactions with the investigatory team. He repeatedly stated his distrust and dislike of governments, corporations, and wealthy people. He said the world is better off without James Salle in it but continually stated he is not responsible for Salle's death.

13. Rabies epidemiology records. An epidemiologist on the investigatory team reviewed information provided by the Xenon Centers for Disease Control concerning recent rabies cases in foxes and humans. There were 6 instances of rabies in foxes in the last 4 years in Smith's hometown and 2 fatal cases in humans, one of whom was Smith's cousin. Both fatal human rabies cases were caused by bites from the same fox, which was eventually captured, euthanized, and tested. Through a public health surveillance effort initiated after the first fatal human case of rabies, public health and animal control officials identified 5 more cases of rabies in foxes, all in animals located on the Western side of the city where Smith lives.

14. Biotech Facility inspection. The facility is run by a non-profit organization that specializes in creating diagnostics for zoonotic pathogens. A reference lab run by the same organization and employing one of Smith's roommates holds 3 stocks of rabies virus. Sequence analysis of all 3 strains indicated none had 2 copies of the G gene. Biotech Facility records indicate it has not conducted experiments with rabies viruses nor sent rabies virus samples to other labs in the last 5 years. Biotech Facility made 2 shipments of rabies virus within the last 10 years, one 8 years ago and another 6 years ago. Notably, both shipments were made to a rabies lab at University Laboratory, and both shipped strains were isolated from wild animals: one from an infected bat and the other from a fox.

15. University Laboratory inspection. University Laboratory is associated with a university in Xenon. It has a robust biomedical research program that includes research with several high consequence pathogens. Smith's other roommate works in a University Laboratory that specializes in flaviviruses and conducts research in mice and non-human primates. The investigatory team tested the flavivirus lab, the animal facility, and the animals for rabies virus, but all samples came back negative. Staff, including Smith's roommate, are interviewed, and while none report suspicious activity, all staff reported experience with viral genetic engineering. The team reviewed several projects conducted by 2 different University Laboratory groups working on rabies. Two of these projects resulted in the enhanced pathogenicity mutations identified in the

samples taken from James Salle, including one project on which Smith's roommate previously worked. Using sequences posted on a publicly available database, the investigatory team compared the rabies sequences published from University Laboratory to the sequences from Salle's samples. The nucleotide similarity between these samples and those from Salle ranged from 86.1% to 96.9% nucleotide similarity.

Specific evidence on Jim Little:

16. Interview with Jim Little. Little and his legal team cooperated with investigators. Little repeatedly denied throwing the dart or genetically modifying rabies. In interviews, Little reported he let Jasper Bill borrow his gym bag in the last year, and Little had gone to theIOM rabies lab where Bill worked to pick it up one day prior to Bill's death. Little acknowledged he knew he was not supposed to enter the rabies lab but said it was an easy place to chat with his friend. Little said he has extensive experience working with CRISPR and viruses.

17. Samples collected from Little's lab. In the IOM pharmacogenetics lab where Little works, the investigatory team used PCR to test surfaces and randomly selected freezer tubes for rabies. No samples were positive for rabies virus.

18. Samples collected from the IOM rabies lab. The team conducted genetic sequencing of rabies samples collected from the IOM rabies lab and compared those sequences to the samples collected from Salle. When a phylogenetic tree was created using these samples, the isolates from the IOM lab grouped together while the isolates from Salle created their own, distinct clade. The highest nucleotide sequence similarity between any sample from the rabies lab and the Salle samples was 83.2%.

19. Genetic engineering signatures in the rabies lab samples. The IOM rabies lab samples contained evidence of genetic engineering, but the virologist on the investigatory team determined all such signatures are consistent with the signatures expected based on the experiments documented in the lab's notebooks. All the samples contained only 1 copy of the G gene, no samples included the enhanced pathogenicity mutations, and none of the genetic manipulation signatures were in the G gene.

20. IOM keycard logs and security footage. The IOM shared 10 years' worth of badge-swipe records from the entire building with the investigatory team. Although Little did not have access to the rabies lab floor, he attempted to access that floor, as well as 2 others to which he did not have access, at least 18 times over the 10-year period. Following the 11 times Little attempted to use his card on the rabies floor,

the badges enabling the first door opening following his attempts belonged to 5 different individuals, with Jasper Bill being responsible for 6 of the openings. Security camera footage from the last year shows Little entering the rabies lab on 2 occasions: one visit lasts 5 minutes, 42 seconds,and the other lasts 25 minutes, 18 seconds. On both occasions, Little entered the lab with Bill. In footage of the second visit, Little left the rabies lab with a bag he did not have when he entered. Little cannot produce the bag when asked to do so.

Conclusion Based on the evidence obtained during the investigation, the investigatory team concluded unanimously that the dart thrown at James Salle on April 5 was purposely contaminated with rabies virus with the intent to harm him. However, the investigatory team is not unanimous in their beliefs about who may have thrown the dart.

A.1.2 Table A1

Pairwise Wilcoxon test with BH correction

Group1	Group2	p_adj	Significance
SurveillanceCCTV	Genomic	0.000	***
Genomic	EventVideo	0.000	***
Witness	Genomic	0.000	***
SurveillanceCCTV	PMXME	0.000	***
SurveillanceCCTV	Metagenomic	0.000	***
Genomic	Environmental	0.000	***
Witness	Metagenomic	0.000	***
Metagenomic	EventVideo	0.000	***
Witness	PMXME	0.000	***
SurveillanceCCTV	PMSR	0.000	***
Metagenomic	Environmental	0.000	***
PMXME	EventVideo	0.000	***
Interviews	Genomic	0.000	***
Metagenomic	Interviews	0.000	***
SurveillanceCCTV	Records	0.000	***
Witness	PMSR	0.000	***
PMXME	Environmental	0.000	***
PMSR	EventVideo	0.000	***
Witness	Records	0.000	***
PMSR	Environmental	0.000	***
Records	EventVideo	0.000	***
PMXME	Interviews	0.000	***
PMXME	Metagenomic	0.000	***
PMSR	Metagenomic	0.000	***
Records	Environmental	0.000	***
PMSR	Interviews	0.000	***
Records	Metagenomic	0.000	***
Records	Genomic	0.000	***
PMXME	Genomic	0.000	***
PMSR	Genomic	0.000	***
Records	Interviews	0.000	***
SurveillanceCCTV	Interviews	0.000	***
SurveillanceCCTV	EventVideo	0.001	***
SurveillanceCCTV	Environmental	0.001	**
Witness	SurveillanceCCTV	0.027	*
Witness	Interviews	0.057	ns
Witness	Environmental	0.159	ns
Records	PMXME	0.176	ns
Witness	EventVideo	0.247	ns
Records	PMSR	0.267	ns
Interviews	EventVideo	0.296	ns

Group1	Group2	p_adj	Significance
Interviews	Environmental	0.494	ns
EventVideo	Environmental	0.664	ns
Metagenomic	Genomic	0.845	ns
PMXME	PMSR	0.845	ns

A.1.3 Table A2

Pairwise Wilcoxon test with BH correction

Group1	Group2	p_adj	Significance
Social media analysis	Genetic Engineering Signatures in the Rabies Lab Samples	0.000	***
Social media analysis	Samples Collected from the IOM Rabies Lab	0.000	***
University Laboratory Inspection	Social media analysis	0.000	***
Social media analysis	Samples Collected from Little's Lab	0.000	***
Genetic Engineering Signatures in the Rabies Lab Samples	Biotech Facility Inspection	0.000	***
Samples Collected from the IOM Rabies Lab	Biotech Facility Inspection	0.000	***
Interview with Roy Smith	Genetic Engineering Signatures in the Rabies Lab Samples	0.000	***
Social media analysis	IOM keycard logs and security footage	0.000	***
Rabies epidemiology records	Genetic Engineering Signatures in the Rabies Lab Samples	0.000	***
University Laboratory Inspection	Biotech Facility Inspection	0.000	***
University Laboratory Inspection	Interview with Roy Smith	0.000	***
Samples Collected from the IOM Rabies Lab	Interview with Roy Smith	0.000	***
Interview with Jim Little	Genetic Engineering Signatures in the Rabies Lab Samples	0.000	***
University Laboratory Inspection	Rabies epidemiology records	0.000	***
University Laboratory Inspection	Interview with Jim Little	0.000	***
Samples Collected from the IOM Rabies Lab	Rabies epidemiology records	0.000	***
Samples Collected from the IOM Rabies Lab	Interview with Jim Little	0.000	***
Samples Collected from Little's Lab	Biotech Facility Inspection	0.000	***
Samples Collected from Little's Lab	Interview with Roy Smith	0.000	***
Social media analysis	Biotech Facility Inspection	0.000	***
Samples Collected from Little's Lab	Interview with Jim Little	0.002	**
Samples Collected from Little's Lab	Rabies epidemiology records	0.002	**
IOM keycard logs and security footage	Interview with Roy Smith	0.002	**
Social media analysis	Interview with Roy Smith	0.002	**
IOM keycard logs and security footage	Interview with Jim Little	0.004	**
Samples Collected from Little's Lab	Genetic Engineering Signatures in the Rabies Lab Samples	0.004	**
IOM keycard logs and security footage	Biotech Facility Inspection	0.005	**
Rabies epidemiology records	IOM keycard logs and security footage	0.005	**
Social media analysis	Rabies epidemiology records	0.008	**
Social media analysis	Interview with Jim Little	0.008	**
University Laboratory Inspection	Samples Collected from Little's Lab	0.040	*

Group1	Group2	p_adj	Significance
IOM keycard logs and security footage	Genetic Engineering Signatures in the Rabies Lab Samples	0.078	ns
Samples Collected from the IOM Rabies Lab	Genetic Engineering Signatures in the Rabies Lab Samples	0.088	ns
University Laboratory Inspection	IOM keycard logs and security footage	0.167	ns
Samples Collected from the IOM Rabies Lab	Samples Collected from Little's Lab	0.167	ns
Interview with Roy Smith	Biotech Facility Inspection	0.346	ns
University Laboratory Inspection	Samples Collected from the IOM Rabies Lab	0.359	ns
Interview with Jim Little	Biotech Facility Inspection	0.362	ns
Samples Collected from the IOM Rabies Lab	IOM keycard logs and security footage	0.452	ns
Rabies epidemiology records	Biotech Facility Inspection	0.719	ns
University Laboratory Inspection	Genetic Engineering Signatures in the Rabies Lab Samples	0.743	ns
Interview with Roy Smith	Interview with Jim Little	0.774	ns
Samples Collected from Little's Lab	IOM keycard logs and security footage	0.798	ns
Rabies epidemiology records	Interview with Jim Little	0.941	ns
Rabies epidemiology records	Interview with Roy Smith	0.996	ns

Table A.1: Summary of test statistics from the Wilcoxon Sign Ranked Test to determine which pairings of evidence have a significant difference in median rank when considering if Salle was deliberately killed with a bioweapon.

Table A.2: Summary of test statistics from the Wilcoxon Sign Ranked Test to determine which pairings of evidence have a significant difference in median rank when considering who threw the dart that hit Salle.

A.2 Scenario 2

Potential Use of Biological Weapons in Sol, Nyx, and Hemera

A.2.1 Text of Scenario

Background Sol, Nyx, and Hemera are adjacent middle- income countries. Significant portions of Sol are controlled by rival cartels.

Year 1: Starting in November of Year 1, health officials in Sol noticed a sharp increase in the number of non-Hodgkin's lymphoma and hepatic cancer cases in the Northern region controlled by the North Sol Cartel. They also noted increases in new cases of birth defects, influenza, and foodborne illness in children and elderly adults living near North Sol Cartel territory.

Year 2: In November of Year 2, health officials in Sol again reported increased case numbers of hepatic cancer and birth defects, as well as severe cases of various infectious diseases. Public health officials suspected dioxin exposure is contributing to the increased disease incidence. Dioxins are carcinogenic chemical compounds generated by certain industrial practices, such as incinerating waste or chemical manufacturing, as well as natural events like volcanic eruptions or forest fires. For most people dietary intake of animal fat is the primary route of exposure to dioxins. An investigation did not reveal significant findings, but rumors circulated that the problem is caused by pesticide production in Nyx, an adjacent country to the north and northwest of Sol.

Year 3: By mid-Year 3, Hemera, an adjacent country to the east of Sol, also reported increased cases of birth defects, severe infectious disease, and hepatic cancers. Hemera's Centers for Disease Control (CDC) initiated a study of the unusual hepatic cancers/tumors and found that the tumor transcriptional signature was similar to scientific studies investigating dioxin exposure. The findings led officials to confirm earlier

rumors that dioxins might be causing the increase in cancer cases. Hemera CDC undertook further studies and found that farmed pigs had elevated dioxin levels. Notably, Nyx reported no increase in incidence of cancers, birth defects, or infectious disease to the World Health Organization (WHO) in Year 3. During a speech, the Deputy Minister of Agriculture in Sol accused Nyx of creating a pesticide bioweapon, leading to bioaccumulation in livestock. Additionally, the Government of Sol accused Nyx of making food dangerous for human consumption, claiming dioxins found in the meat caused cancer and other adverse events. However, Hemera leaders believed Sol's accusations were premature. Instead, Hemera CDC expanded its study to survey dioxin levels in both wild and farmed pigs. They found that bioaccumulation of dioxin was affecting about 90% of farmed pigs but only 2% of wild pigs. Once the potential link to pigs was established, public health authorities in the region worked with the World Organisation for Animal Health (formerly the OIE), the Food and Agriculture Organization (FAO), and the WHO to investigate potential associations between elevated dioxin levels in pigs and increased incidence of human diseases. Part of the work included assessing past swine disease outbreaks in the region. The collaborative team found there was an outbreak of African Swine Fever (ASF) in the region about 10 years prior to Year 1. Since then, intense efforts to develop and produce ASF vaccines for pigs have been underway. One vaccine was developed and produced in a facility in Hemera, near the borders of Sol and Nyx. This vaccine was given to farmed pigs in the region for 4 years prior to Year 1, with one dose given at age 6 weeks, another dose at age 6 months, and a final dose at age 1 year. Investigators assessed the original formula of the vaccine and did not find any components that could lead to high dioxin concentrations in pigs.

Year 4: Sol continued to maintain that a biological weapon was used. Nyx maintained they are not to blame for the increased dioxins in swine and strongly refuted allegations they created a biological weapon. Public sentiment in the region, including in Hemera, shifted to widely support the biological weapon theory, though there was widespread disagreement over who might be the responsible party.

Sol and Hemera got Nyx to agree to partner to investigate potential bioweapon use. They initiated an investigation into the increased cases of human disease and increased dioxin bioaccumulation in pigs. During the investigation, Hemera and Sol culled their farmed pigs. Nyx does not farm pigs and therefore had none to cull. The South Sol Cartel did not announce a farmed pig cull, but the cartel abruptly stopped swine sales. The North Sol Cartel did not cull their pigs and continued swine sales, though demand was

drastically reduced.

Investigation Overview The investigatory team used the procedures set forth in the United Nations Guidelines and Procedures for the timely and efficient investigation of reports of the possible use of chemical and bacteriological (biological) or toxin weapons (A/44/561) to guide its investigation. The team collected a wide range of evidence, some on-site and some off-site. While the investigatory team had fairly open access to Hemera and Nyx, and the government-controlled 2 parts of Sol, they were not able to enter the cartel-controlled areas of Sol. Therefore, they were not able to collect samples in Sol's cartel territories and instead relied on information provided by the cartels, samples from nongovernmental organizations (NGOs) working in the area, and satellite imagery.

Evidence

1. Environmental sampling for dioxins. The investigatory team sampled rivers, lakes, wetlands, and the ocean, indoor and outdoor air, and soil across the 3-nation region. Those samples did not have significantly elevated levels of dioxins or its derivatives compared to environmental samples taken 10 years prior by national authorities in these countries or levels reported by academic groups 8 years prior.

2. Epidemiological data from national authorities. The investigatory team collected epidemiological data from NGOs working in the area, as well as national and local public health agencies. The epidemiologists on the investigatory team analyzed this data and concluded there was an increase in new cases of cancers, birth defects, and certain infectious diseases in the first half of Year 1 in the northern region of Sol. The epidemiologists also found increases in new cases of cancers, birth defects, and infectious diseases in the rest of Sol in Year 2, followed by increases in the same conditions in Year 3 in Nyx and Hemera. The epidemiologists noted that the same types of cancers (primarily non-Hodgkin's lymphoma and hepatic cancers), birth defects, and infectious diseases were elevated in each of the regions.

3. Medical records from patients in Hemera, Nyx, and Sol. The investigatory team obtained anonymized medical records from the NGOs in Northern Sol that treated patients in Years 1 and 2. The medical professionals on the investigatory team reviewed these files and found there were elevated numbers of new cases of non-Hodgkin's lymphoma or hepatic cancers, new cases of foodborne and flu-like illnesses, and an increased number of birth defects in people in the northern region of Sol in first half of Year 1 and in the other areas of Sol in Year 2. The medical professionals on the investigatory team determined the

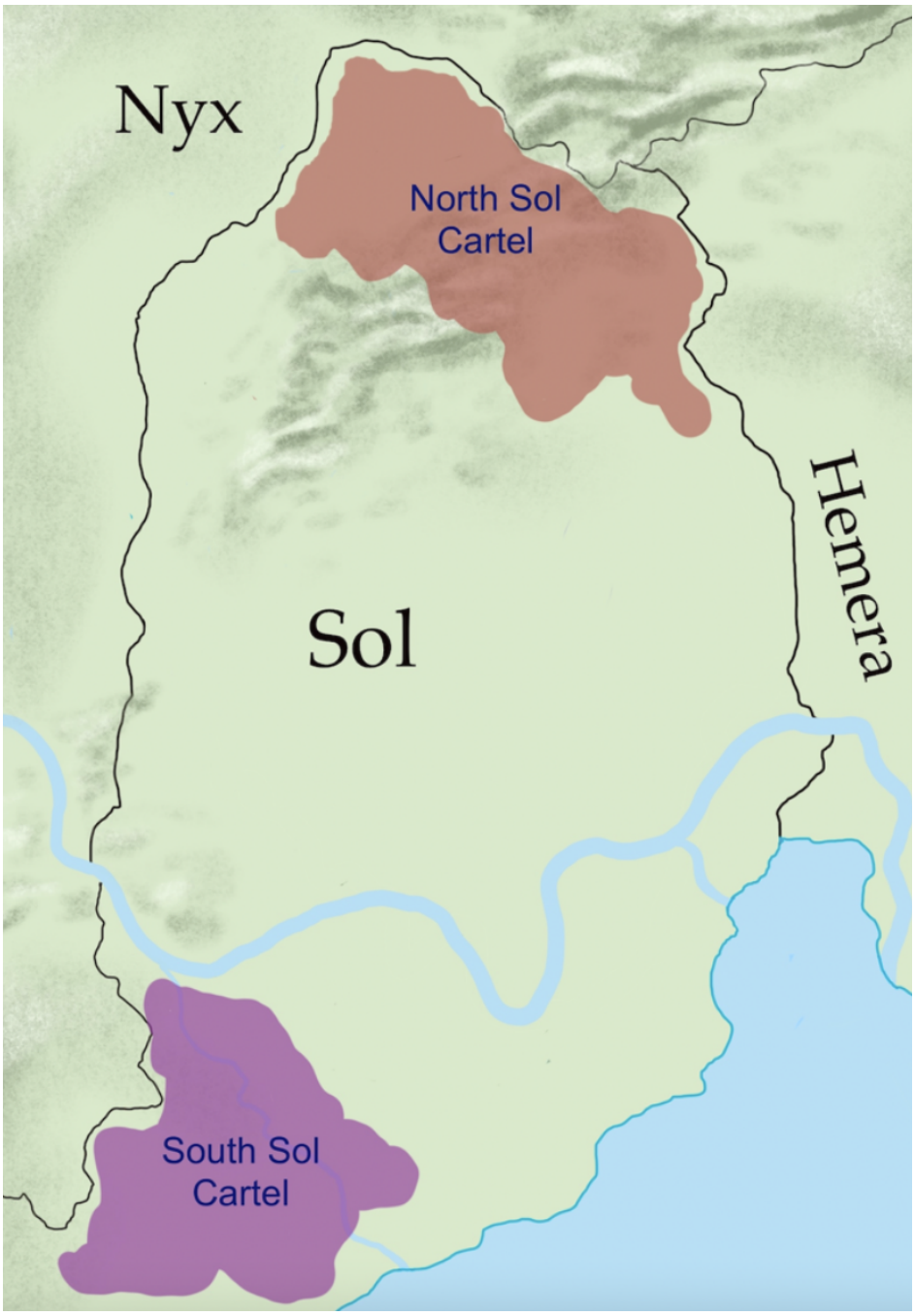

Figure A.1: Map of Sol, Nyx, and Hemera for Scenario 2

recorded cases were diagnosed appropriately.

The investigatory team also obtained anonymized medical records from hospitals in Hemera, Nyx, and government-controlled areas of Sol that treated patients in Years 1 through 4. The team's medical professionals reviewed these files and found there were also elevated numbers of new cases of non-Hodgkin's lymphoma or hepatic cancers, new cases of foodborne and flu-like illnesses, and birth defects in patients from government-controlled areas of Sol starting in Year 2, and in patients from Hemera and Nyx starting in Year 3. The investigators noted that while Sol and Hemera had hundreds of cases more per year in Years 3 and 4 compared to Years 1 and 2, Nyx only had about 10 more cases in Years 3 and 4 compared to Years 1 and 2.

4. Transcriptomic analysis of samples from current patients. The investigatory team collected samples from patients currently experiencing non-Hodgkin's lymphoma or hepatic cancers in government controlled areas of Sol, Nyx, and Hemera. Samples were also collected from patients experiencing non-Hodgkin's lymphoma or hepatic cancers in other parts of the world. RNA sequencing was used to assess the transcriptomes of the patients in a transcriptome-wide association study. The investigatory team found transcription levels of 14 genes significantly elevated in at least 90% of the Sol, Nyx, and Hemera patients but in fewer than 10% of patients from other countries. In some cases, transcription profiles can indicate if different tumor developed from a common exposure source. The geneticists and epidemiologists on the investigatory team conducted a literature review and found that 9 of these genes were linked to dioxin exposure in prior research studies.

5. Surveying pigs for dioxin bioaccumulation. The investigatory team conducted a survey of wild and farmed pigs in Hemera, Nyx, and government-controlled areas of Sol where possible. Nyx reported having no farmed pigs and would not provide access to farms for the investigatory team to verify this information. Results of the survey in pigs found elevated levels of dioxin in 85% to 91% of farmed pigs over 6 months old in Hemera and government-controlled areas of Sol and in 2% to 4% of wild pigs in Hemera, Nyx, and government-controlled areas of Sol.

6. Satellite imagery analysis of chemical processing plants and other facilities in Nyx. The investigatory team gathered satellite imagery from Years 0 to 4 of the areas around herbicide and chemical manufacturing research and development facilities in Nyx. Based on the satellite images, there were no

changes in activity level or facility growth during this time. No evidence of dumping at the facilities or their immediate surrounding areas is apparent in the satellite images. Additionally, no pigs (wild or farmed) appeared to be in the vicinity of any of these facilities at any time. The animals and habitats surrounding the facility did not have significant changes across the period observed.

7. Satellite images of cartel territories in Sol. The investigatory team gathered satellite imagery of the cartel-controlled areas of Sol from Years 0 to 4. Wild and farmed pigs were seen in each of the territories. Investigators noted that while each cartel had pig farms, the North Sol cartel had abnormally stringent practices to keep their cartel- owned pigs away from all other pigs in the area. Images also suggested that cartel members only slaughtered and processed pigs from their own farms. Additionally, the healthcare facilities, mortuaries, and funeral homes in the North Sol cartel territory did not appear to have an increase in visitors or activity other than in early Year 1, whereas the same facilities in the South Sol cartel territory did have sustained increases from Year 2 onward.

8. Previously published research article. The investigatory team uncovered a research paper published by a lab in Hemera that designed bacterial plasmids that could break down environmental dioxins. In the process of creating the plasmids that break down dioxins, the lab also created plasmids that contained a series of genes that when expressed, created a synthesis pathway for dioxins. The authors of the paper were from Hemera and Sol, but when the team interviewed them and investigated their backgrounds, no suspicious behavior or activities were found.

9. Interviews with local veterinarians. The investigatory team interviewed veterinary professionals in Sol, Nyx, and Hemera. Veterinarians reported they had been giving the ASF vaccine per protocol since it became available to all farms in Hemera and government-controlled areas of Sol. They also said they had started seeing signs of disease in pigs starting in late Year 1 in Sol and Year 2 in Hemera. The veterinarians said they did not know of any pig farms in Nyx, but all the cartel-controlled areas of Sol contained farms with pigs. The investigatory team secretly interviewed 2 veterinarians who worked on farms in the South Sol cartel regions. Both reported using the ASF vaccine on a normal schedule, but they had seen disease in the pigs on their farms in Year 1.

10. Intelligence signals in the region. The investigatory team reached out to intelligence agencies in the region for briefings related to the cartels, environment, pigs, health, and bioweapons. There were intelligence

signals suggesting the cartels were looking to expand territory and gain money. There was no intelligence suggesting Nyx or any other state was interested in using a bioweapon or intentionally trying to pollute the environment. However, there was evidence that members of the North Sol cartel were in regular contact with individuals living in Hemera, including individuals working at the facility that produces the ASF vaccine.

A.2.2 Table A3

Pairwise Wilcoxon test with BH correction

Group1	Group2	p_adj	Significance
Surveying_Pigs	Satellite_ChemProcessing	0.000	***
Surveying_Pigs	Previous_Research	0.000	***
Surveying_Pigs	Interview_Vets	0.002	**
Satellite_ChemProcessing	Environmental_Sampling	0.006	**
Previous_Research	Environmental_Sampling	0.007	**
Surveying_Pigs	Epi_Data	0.007	**
Surveying_Pigs	Intel_Signals	0.007	**
Surveying_Pigs	Medical_Records	0.009	**
Transcriptomic	Satellite_ChemProcessing	0.012	*
Surveying_Pigs	Satellite_Sol	0.012	*
Transcriptomic	Previous_Research	0.013	*
Previous_Research	Medical_Records	0.015	*
Satellite_Sol	Previous_Research	0.017	*
Satellite_ChemProcessing	Medical_Records	0.019	*
Satellite_Sol	Satellite_ChemProcessing	0.019	*
Transcriptomic	Surveying_Pigs	0.019	*
Surveying_Pigs	Environmental_Sampling	0.033	*
Previous_Research	Intel_Signals	0.056	ns
Previous_Research	Epi_Data	0.066	ns
Satellite_ChemProcessing	Intel_Signals	0.082	ns
Previous_Research	Interview_Vets	0.095	ns
Satellite_ChemProcessing	Epi_Data	0.110	ns
Interview_Vets	Environmental_Sampling	0.132	ns
Satellite_ChemProcessing	Interview_Vets	0.138	ns
Transcriptomic	Interview_Vets	0.179	ns
Epi_Data	Environmental_Sampling	0.357	ns
Intel_Signals	Environmental_Sampling	0.434	ns
Medical_Records	Interview_Vets	0.434	ns
Satellite_Sol	Interview_Vets	0.493	ns
Transcriptomic	Epi_Data	0.525	ns
Transcriptomic	Intel_Signals	0.553	ns
Satellite_Sol	Environmental_Sampling	0.575	ns
Satellite_ChemProcessing	Previous_Research	0.575	ns
Medical_Records	Environmental_Sampling	0.659	ns
Satellite_Sol	Epi_Data	0.659	ns
Medical_Records	Epi_Data	0.683	ns
Transcriptomic	Satellite_Sol	0.683	ns
Interview_Vets	Epi_Data	0.730	ns
Interview_Vets	Intel_Signals	0.730	ns
Transcriptomic	Medical_Records	0.730	ns
Medical_Records	Intel_Signals	0.732	ns

Group1	Group2	p_adj	Significance
Satellite_Sol	Intel_Signals	0.760	ns
Transcriptomic	Environmental_Sampling	0.932	ns
Intel_Signals	Epi_Data	0.932	ns
Satellite_Sol	Medical_Records	0.942	ns

A.2.3 Table A4

Pairwise Wilcoxon test with BH correction

Group1	Group2	p_adj	Significance
Surveying_Pigs	Previous_Research	0.000	***
Surveying_Pigs	Satellite_ChemProcessing	0.001	***
Surveying_Pigs	Medical_Records	0.001	**
Surveying_Pigs	Epi_Data	0.003	**
Surveying_Pigs	Interview_Vets	0.003	**
Satellite_Sol	Previous_Research	0.003	**
Transcriptomic	Surveying_Pigs	0.003	**
Surveying_Pigs	Environmental_Sampling	0.005	**
Previous_Research	Intel_Signals	0.005	**
Previous_Research	Environmental_Sampling	0.014	*
Previous_Research	Epi_Data	0.016	*
Transcriptomic	Previous_Research	0.025	*
Satellite_Sol	Satellite_ChemProcessing	0.032	*
Previous_Research	Interview_Vets	0.035	*
Satellite_ChemProcessing	Intel_Signals	0.055	ns
Previous_Research	Medical_Records	0.075	ns
Satellite_Sol	Medical_Records	0.097	ns
Satellite_ChemProcessing	Previous_Research	0.106	ns
Satellite_ChemProcessing	Epi_Data	0.136	ns
Medical_Records	Intel_Signals	0.139	ns
Satellite_ChemProcessing	Environmental_Sampling	0.150	ns
Surveying_Pigs	Intel_Signals	0.168	ns
Satellite_Sol	Interview_Vets	0.182	ns
Transcriptomic	Satellite_Sol	0.208	ns
Interview_Vets	Intel_Signals	0.240	ns
Transcriptomic	Intel_Signals	0.299	ns
Satellite_Sol	Environmental_Sampling	0.324	ns
Transcriptomic	Satellite_ChemProcessing	0.324	ns
Surveying_Pigs	Satellite_Sol	0.330	ns
Satellite_Sol	Epi_Data	0.336	ns
Satellite_ChemProcessing	Interview_Vets	0.376	ns
Intel_Signals	Epi_Data	0.388	ns
Intel_Signals	Environmental_Sampling	0.388	ns
Medical_Records	Environmental_Sampling	0.388	ns
Medical_Records	Epi_Data	0.399	ns
Interview_Vets	Epi_Data	0.537	ns
Satellite_ChemProcessing	Medical_Records	0.639	ns
Transcriptomic	Medical_Records	0.653	ns
Interview_Vets	Environmental_Sampling	0.678	ns
Transcriptomic	Environmental_Sampling	0.684	ns
Transcriptomic	Epi_Data	0.700	ns

Group1	Group2	p_adj	Significance
Satellite_Sol	Intel_Signals	0.757	ns
Medical_Records	Interview_Vets	0.757	ns
Transcriptomic	Interview_Vets	0.920	ns
Epi_Data	Environmental_Sampling	0.977	ns

A.3 Survey

Assessing Evidence from an Investigation of Biological Events

Survey from the Johns Hopkins Center for Health Security

Consent

We are conducting a research study about perceptions of evidence for attribution of biological agents. You have been sent two fictious scenarios describing a possible biological event. Each scenario includes background information, a description of the event, and provides a list of evidence gathered when investigating the event. Each piece of evidence is briefly described. This survey will ask questions about the evidence and what you believe happened in the scenario, followed by brief demographics questions.

We expect the activity to take approximately 2 hours total to complete. All survey responses will be on a not-for-attribution basis. Responding to this survey request is voluntary; it is your choice. You may choose not to answer any question that we ask. Completing this survey and submitting it to us means you consent to participate in the study.

Instructions

Please refer to the PDFs attached to the email that provided you with the link to this survey.

For this exercise, you are an advisor to a decision maker who has ordered an investigation into a biological event. An investigatory team has conducted the investigation and provided the decision maker with a report of their findings (summarized in the PDFs attached to the email). The decision maker has asked you to review the results of the investigation in order to give advice concerning next steps. Please answer the questions in this survey as this decision maker's advisor.

You have been given two scenarios. The first scenario focuses on the death of James Salle. The second scenario assess a potential bioweapon use in Sol, Nyx, and Hemera. For both scenarios, you will be asked to rank the evidence reported by the investigatory team by relative influence on drawing your conclusion. You will also be asked to answer questions regarding your conclusions of the investigation and your opinions on the evidence.

Please refer back to the PDFs provided in the email for these scenarios as needed while completing this survey.

Scenario 1: Death of James Salle

Please refer to the PDF titled "Scenario 1" for the next three sections.

Scenario 1: Was this a deliberate biological attack?

Please refer to the Scenario 1 PDF attached to the email.

1. Please rank the following evidence in order of **most influential** for determining **if a biological weapon was used**.

- Event venue surveillance video recording.
- Surveillance footage from Argon Capital City CCTV.
- Environmental sampling of venue.
- Witness statements.
- Medical records from the hospital where Salle died.
- Post- mortem samples collected from Xenon Medical Examiner.
- Post-mortem samples reanalysis.
- Genomic analysis of post-mortem samples.
- Metagenomic sequencing of swabs from dart.
- Interviews with Salle's personal staff and close contacts.

2. Is there enough evidence included in this report for you to conclude James Salle died from a deliberate attack using a biological weapon? Please explain why or why not.

3. How confident are you that this was a deliberate biological attack?

0	1	2	3	4	5	6	7	8	9	10

Not at all confident Extremely confident

4. How credible and/or trustworthy did you find the evidence included in this section and why?

5. If you could request one additional piece of evidence in order to determine if this was a deliberate attack using a biological weapon, what would that piece of evidence be?

Scenario 1: Who threw the dart?

Please refer to the Scenario 1 PDF attached to the email.

6. Who do you believe is responsible for throwing the dart?

- ◯ Roy Smith
- ◯ Jim Little
- ◯ Another person
- ◯ No one threw a dart

7. Please explain why you selected the answer you chose for Question 5 'Who threw the dart?'.

8. How confident are you that you have identified the correct suspect?

0	1	2	3	4	5	6	7	8	9	10

Not at all confident Extremely confident

9. Please rank the following evidence in order of **most influential** for determining **who threw the dart**.

- Social Media Analysis
- Interview with Roy Smith.
- Rabies epidemiology records.
- Biotech Facility inspection.
- University Laboratory inspection.
- Interview with Jim Little.
- Samples collected from Little's Lab.
- Samples collected from the IOM rabies lab.
- Genetic engineering signatures in the rabies lab’s samples.
- IOM keycard logs and security footage.

10. Which evidence, if any, did you find exonerating for a suspect and why?

11. How credible and/or trustworthy did you find the evidence included in this section and why?

12. If you could request one additional piece of evidence in order to determine who threw the dart, what would that piece of evidence be and why?

Scenario 1: Overarching questions.

13. Was there any evidence that you found difficult to assess or incorporate into your thinking about this scenario?

14. How did you assess credibility when considering the evidence? What factors or characteristics contributed to your conclusion(s) concerning credibility?

15. Are there any types of evidence you would have liked to see but were not included in the scenario? Was there any evidence included that you found inappropriate for the scenario?

16. Do you have any additional thoughts related to evidence in Scenario 1 you would like to share?

Scenario 2: Possible Bioweapon in Nyx, Sol, and Hemera.

Please refer to the PDF titled "Scenario 2" for the following 3 sections.

Scenario 2: What type of event is this?

Please refer to the Scenario 2 PDF attached to the email.

17. Do you believe any of the trends in disease described in the scenario are due to non-natural events?

- ◯ Yes- at least some of the cases of disease are associated with a deliberate or accidental event.
- ◯ No- none of the cases of disease are associated with a deliberate or accidental event.

18. Why do you believe the event is naturally occurring or non-naturally occurring?

19. Do you believe this is accidental or deliberate? (skip if you believe naturally occurring)

- ◯ Accidental
- ◯ Deliberate

20. Why do you believe the event is deliberate or accidental in nature? What evidence did you use to come to your conclusion? (skip if you believe naturally occurring)

21. How confident are you in your assessment of the origins of this event?

0	1	2	3	4	5	6	7	8	9	10

Not at all confident Extremely confident

22. Please rank the following evidence in order of **most influential** for determining **if an intentional biological event has occurred**.

- Environmental sampling for dioxins.
- Epidemiological data from national authorities.
- Medical records from patients in Hemera, Nyx, and Sol.
- Transcriptomic analysis of samples from current patients.
- Surveying pigs for dioxin bioaccumulation.
- Satellite imagery analysis of chemical processing plants and other facilities in Nyx.
- Satellite images of Sol's cartels' territories.
- Previously published research article.
- Interviews with local veterinarians.
- Intelligence signals in the region

23. If you could request one additional piece of evidence in order to determine what type of event is occurring, what would it be and why?

Scenario 2: Who is responsible?

Please refer to the Scenario 2 PDF attached to the email.

24. Who or what do you believe is responsible for this event and why?

25. Please rank the following evidence in order of **most influential** for determining **who or what is responsible for the biological event**.

- Environmental sampling for dioxins.
- Epidemiological data from national authorities.
- Medical records from patients in Hemera, Nyx, and Sol.
- Transcriptomic analysis of samples from current patients.
- Surveying pigs for dioxin bioaccumulation.
- Satellite imagery analysis of chemical processing plants and other facilities in Nyx.
- Satellite images of Sol's cartels' territories.
- Previously published research article.
- Interviews with local veterinarians.
- Intelligence signals in the region

26. How confident are you in your assessment of who or what is responsible for this event?

0	1	2	3	4	5	6	7	8	9	10

Not at all confident Extremely confident

27. If you could request one additional piece of evidence in order to determine who or what is responsible for this event, what would it be and why?

Scenario 2: Overarching questions.

28. Was there any evidence that you found difficult to assess or incorporate into your thinking about this scenario?

29. How credible and/or trustworthy did you find the evidence included in Scenario 2 and why?

30. How did you assess credibility when considering the evidence? What factors or characteristics contributed to your conclusion(s) concerning credibility?

31. Are there any types of evidence you would have liked to see but were not included in the scenario? Was there any evidence included that you found inappropriate for the scenario?

32. Do you have any additional thoughts related to evidence in Scenario 2 you would like to share?

Demographics

33. Which field(s) reflects your area of expertise?

- [ ] Evolutionary and/or Molecular Biology
- [ ] Biotechnology
- [ ] Microbiology (bacteriology, virology, etc.)
- [ ] Social Science
- [ ] Public Health
- [ ] International Relations and/or Diplomacy
- [ ] Law
- [ ] Forensic Sciences
- [ ] Biosecurity
- [ ] Biosafety
- [ ] Chemical or Nuclear Security
- [ ] Nonproliferation
- [ ] Food Safety
- [ ] Environmental Sciences
- [ ] Other (please specify in next question)

34. If you selected "Other" in the previous question, please name the field(s) you feel best reflect your area of expertise.

35. How many years of experience do you have in your field?

- ◯ 0 to 5
- ◯ 6 to 10
- ◯ 11 to 20
- ◯ 21 to 30
- ◯ 31 to 40
- ◯ 41 to 50
- ◯ Over 50 years

36. What is your gender?

- ◯ Woman
- ◯ Man
- ◯ Non-binary
- ◯ Other
- ◯ Prefer not to say

37. What is your nationality?

Thank you for filling out this survey!

We look forward to sharing results from this work in the coming months.

- Lane Warmbrod and Gigi Gronvall

Table A.3: Summary of test statistics from the Wilcoxon Sign Ranked Test to determine which pairings of evidence have a significant difference in median rank when considering the nature of the biological event in Scenario 2.

Table A.4: Summary of test statistics from the Wilcoxon Sign Ranked Test to determine which pairings of evidence have a significant difference in median rank when who may be responsible for the biological event in Scenario 2.