safety

MDPI

Article

Enhancing Laboratory Resilience: Development and Expert Validation of Risk-Based Emergency Drill Scenarios for BSL-2/ABSL-2 Facilities

Shinhao Yang [1,2,*], Hsiao-Lin Huang [3], Pei-Ling Kuo [2], Yu-Chin Chiang [2] and Yen-An Chen [2]

[1] Environmental Sustainability Lab, Natural Science, Center for General Education, CTBC Business School, No. 600, Section 3, Taijiang Boulevard, Annan District, Tainan City 709, Taiwan
[2] AI Environmental Engineering and Circular Economy Research Center, CTBC Business School, No. 600, Section 3, Taijiang Boulevard, Annan District, Tainan City 709, Taiwan; paney1218@gmail.com (P.-L.K.); biosafetytaiwan@gmail.com (Y.-C.C.); yananchen1658@gmail.com (Y.-A.C.)
[3] Department of Occupational Safety and Health, Chia Nan University of Pharmacy and Science, No. 60, Section 1, Erren Road, Rende District, Tainan City 717, Taiwan; hlhuang@mail.cnu.edu.tw
* Correspondence: shinhaoyang@ntu.edu.tw; Tel.: +886-6-287-3198

Abstract

This study develops and validates risk-based emergency response scenarios for Biosafety Level 2 (BSL-2) and Animal Biosafety Level 2 (ABSL-2) facilities. Utilizing Bow-tie analysis, three multidimensional scenarios were constructed: infrastructure failure, biosecurity breach, and compound disaster. Four domain experts independently evaluated the scripts using the Content Validity Index (CVI), with an absolute consensus threshold of I-CVI = 1.00. To address operational gaps identified during initial evaluations, the revised protocols were strictly aligned with the Taiwan Centers for Disease Control (CDC) mandatory reporting thresholds for high-hazard incidents. Furthermore, the scripts explicitly defined the Incident Command System (ICS) to prevent communication fragmentation and integrated the NC3Rs tunnel handling technique to minimize occupational bite risks. Following these targeted refinements, all items achieved absolute expert consensus. This research translates static biosafety regulations into dynamic, stress-tested training tools. By providing a standardized instrument for resilience assessment, this study equips frontline personnel with the critical capacity to navigate cascading crises while strictly adhering to a "life safety first" paradigm.

Keywords: biosafety; emergency drill; Content Validity Index (CVI); Bow-tie analysis; NC3Rs

Check for updates

Academic Editor: Raphael Grzebieta

Received: 7 April 2026
Revised: 10 June 2026
Accepted: 12 June 2026
Published: 18 June 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Biosafety Level 2 (BSL-2) and Animal BSL-2 (ABSL-2) laboratories serve as foundational infrastructure for global public health defense, infectious disease diagnostics, and biomedical research [1]. Given the routine handling of high-consequence pathogens, these facilities require stringent management protocols. Within this context, distinguishing between biosafety and biosecurity is essential [2]. Biosafety focuses on protecting laboratory personnel from unintended occupational exposure to biological agents or animal hazards. In contrast, biosecurity involves preventing the escape or unauthorized release of pathogens and infected animals, thereby mitigating risks to external populations. While existing protocols primarily emphasize biosafety, practical tools to evaluate dynamic biosecurity vulnerabilities remain limited. The risk-based scenario framework developed in this

study provides an evaluation tool applicable to biosecurity assessments, particularly in ABSL-2 and ABSL-3 animal facilities. Historically, laboratory safety management has heavily relied on static standard operating procedures (SOPs) and physical engineering controls to ensure regulatory compliance [1,2]. However, modern safety science indicates that to effectively navigate unpredictable threats, institutions must transition into High Reliability Organizations (HROs) [3,4]. HROs emphasize systemic "resilience"—the capacity to absorb, adapt to, and rapidly recover from unexpected disruptions and cascading failures [5]. In Taiwan, the Centers for Disease Control (Taiwan CDC) enforces strict regulations through the Regulations Governing Management of Infectious Biological Materials, mandating that all certified biological laboratories conduct regular emergency response drills to ensure institutional preparedness [6,7]. While physical biocontainment technologies have advanced significantly, embedding organizational resilience into the behavioral reflexes of frontline personnel remains a critical challenge for laboratory administrators [3].

Compared to maximum containment facilities, BSL-2 and ABSL-2 laboratories experience a significantly higher frequency of routine operations. Prolonged exposure to these low-to-moderate risk environments frequently breeds the "normalization of deviance" [8,9]. As highlighted by Reason's system approach to human error, when minor instances of non-compliance—such as bypassing standard biocontainment steps—rarely result in immediate adverse events, personnel subconsciously degrade their adherence to safety protocols over time [10]. This phenomenon of rationalizing abnormal behaviors directly contradicts the core principles of HROs and serves as a latent condition that provokes severe accidents [8–10].

Furthermore, real-world laboratory crises seldom manifest as isolated biological spills. Recent scientific research highlights that modern crises increasingly occur as "compound disasters" [11]. A compound disaster involves multiple independent hazards—such as infrastructure failures, extreme environmental events, and biological exposures—overlapping in space and time, thereby generating non-linear and unpredictable risks [11]. Under extreme time pressure and high cognitive load, human error probabilities surge, rendering static, linear SOPs inadequate [10]. This rigidity triggers severe priority conflicts for frontline personnel (e.g., deciding whether to prioritize immediate life-safety evacuation during an earthquake or to maintain strict biocontainment to prevent pathogen dissemination). Consequently, rigorously evaluating system performance through realistic preparedness exercises is crucial, and current emergency management systems urgently need to integrate dynamic risk frameworks to replace inflexible response paradigms [12–14].

Emergency preparedness exercises and simulation drills are universally recognized as essential mechanisms for validating organizational resilience, identifying systemic vulnerabilities, and enhancing interdisciplinary coordination before an actual crisis occurs [13,15]. However, a significant gap exists between the design of these exercises and their practical effectiveness. As systemic reviews of public health and healthcare emergency exercises reveal, many drills suffer from arbitrary evaluation methods and are reduced to "check-the-box" administrative formalities designed merely to satisfy regulatory mandates, rather than to rigorously stress-test response capabilities [13,16].

While various theoretical risk assessment models have been proposed to quantify occupational hazards, there is a distinct lack of methodological research on translating these theoretical data into dynamic, actionable "drill scenarios" [17]. For instance, although the Taiwan Biosafety Association recently drafted a comprehensive template for BSL-2/ABSL-2 emergency drill plans, a standardized framework to integrate these generic templates with institution-specific, dynamic risk assessments is still lacking. Furthermore, currently available simulation materials often lack an integrated systems perspective, failing to effectively challenge and measure response protocols under extreme stress [16]. This disconnect

between theoretical safety governance and practical drill execution leaves hidden vulnerabilities in laboratory biocontainment largely unaddressed until a real disaster strikes, highlighting the urgent need for validated, risk-based scenario development tools [15,18].

To bridge the aforementioned research gaps, this study aims to develop and systematically validate a "Risk-Based Emergency Drill Scenario Framework" specifically tailored for BSL-2 and ABSL-2 facilities. Moving beyond generic, single-hazard templates, this study constructed three high-impact, multi-dimensional drill scenarios designed to rigorously stress-test organizational boundaries and resolve priority conflicts: (1) a facility infrastructure failure (a severe ceiling leak compromising a biological safety cabinet); (2) a critical biosecurity breach (an unrecovered escaped infected animal); and (3) a compound disaster (an earthquake occurring concurrently with an animal escape).

To ensure the scientific rigor, operational realism, and practical applicability of these designed scenarios, this study employed the Content Validity Index (CVI) for expert validation. CVI is a robust quantitative methodology extensively utilized to establish the validity of nursing assessments, emergency medical protocols, and disaster preparedness instruments [19,20]. By applying CVI to evaluate laboratory drill scenarios, this study transforms static safety compliance into a dynamic, evidence-based resilience evaluation tool. Ultimately, this validated framework provides global biological laboratories and public health institutions with a scientifically sound strategy to proactively enhance their biosafety, biosecurity, and comprehensive emergency management capabilities.

2. Materials and Methods

2.1. Study Design

This study employs a structured, two-phase methodological framework to develop, evaluate, and systematically validate risk-based emergency drill scenarios tailored for Biosafety Level 2 (BSL-2) and Animal BSL-2 (ABSL-2) facilities. The overarching goal is to transform static safety regulations into dynamic resilience stress tests that effectively expose latent vulnerabilities within the laboratory system. The overall research flowchart is illustrated in Figure 1.

- Phase 1: Risk-Based Scenario Development via Bow-tie Analysis

The first phase focuses on scenario conceptualization and structural design. Utilizing the Bow-tie methodology, this phase maps the complex pathways between specific hazards, top events, and potential consequences. By identifying critical preventive and reactive barriers, the study systematically constructed three high-impact, multi-dimensional drill scenarios. These scenarios were specifically designed to transcend routine single-hazard events and deliberately induce priority conflicts among personnel. The three selected scenarios encompass: (1) a facility infrastructure failure (a biological safety cabinet compromised by severe ceiling water leakage), (2) a biosecurity breach (an unrecovered escaped infected animal), and (3) a compound disaster (an earthquake triggering simultaneous infrastructure damage and animal escape).

- Phase 2: Quantitative Expert Validation Using the Content Validity Index (CVI)

The second phase aims to rigorously validate the scientific accuracy, operational realism, and practical applicability of the developed scenarios. A structured evaluation instrument was drafted, and four domain experts was invited to assess the scenarios across four core dimensions: realism, risk relevance, procedural accuracy, and structural completeness. The Content Validity Index (CVI) was subsequently employed as the quantitative metric to evaluate expert consensus. Given that exactly four experts were consulted, strict threshold criteria were applied to ensure that the final scenario framework meets the highest academic and practical standards for laboratory emergency preparedness.

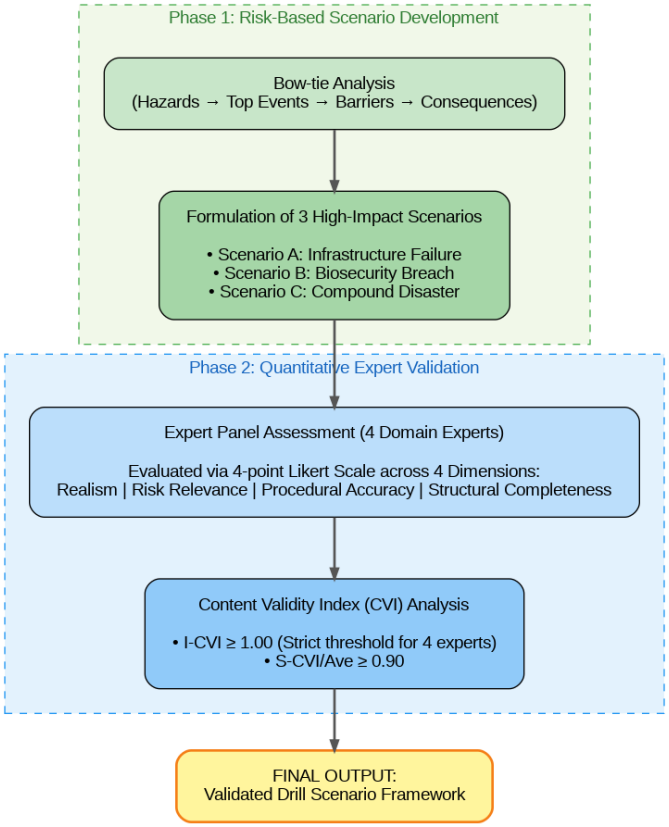


Figure 1. The two-phase methodological framework for developing and validating risk-based emergency drill scenarios in BSL-2/ABSL-2 facilities.

2.2. Phase 1: Risk-Based Scenario Development via Bow-Tie Analysis

2.2.1. The Bow-Tie Methodology

To systematically translate static safety assessments into dynamic drill scenarios, this study employed the Bow-tie methodology. The Bow-tie model is a highly visual and structured risk evaluation tool that integrates the principles of Fault Tree Analysis (identifying causes) and Event Tree Analysis (identifying consequences). In recent years, it has been increasingly adopted in complex socio-technical systems, including healthcare and hazardous industries, for its robust capability to visually map and manage compounding risk pathways [21]. At the center of the diagram is the "Top Event," which represents the moment control over a hazard is lost. The left side of the Bow-tie delineates the potential threats (causes) that can lead to the Top Event and the "Preventive Barriers" designed to intercept them. The right side outlines the potential consequences and the "Reactive Barriers" (mitigation strategies) required to minimize harm.

A critical component of the Bow-tie model utilized in this study is the identification of "Escalation Factors" (also known as barrier decay mechanisms). Escalation factors are secondary conditions or events—such as power outages, extreme weather, or simultaneous emergencies—that compromise or entirely disable a designated safety barrier. By intentionally embedding these escalation factors into the Bow-tie framework, this study generated stress-test scenarios that force personnel to navigate compound risks rather than straightforward, single-variable accidents.

2.2.2. Scenario Formulation

Based on the Bow-tie framework and the specific operational guidelines for BSL-2 and ABSL-2 facilities, three high-impact, multi-dimensional drill scenarios were formu-

lated. Each scenario targets specific barrier failures and assesses different dimensions of organizational resilience:

- Scenario A: Infrastructure Failure (Facility Leakage Compromising BSC).

Concept: This scenario simulates an extreme weather event causing a facility failure that impacts primary engineering controls.

Parameters: During routine operations, heavy rainfall causes a sudden ceiling leak directly above a functioning Biological Safety Cabinet (BSC). The water intrusion potentially compromises the HEPA filter's integrity and leads to aerosol dissemination.

Evaluation Focus: Tests the personnel's ability to immediately mitigate electrical hazards (shutting down BSC power and pulling down the sash), initiate evacuation, enforce a 30-min aerosol settling period, and safely execute full-room surface decontamination.

- Scenario B: Critical Biosecurity Breach (Unrecovered Escaped Infected Animal).

Concept: This scenario focuses on a dynamic biosecurity failure where an infected animal escapes the primary enclosure and remains unaccounted for, presenting an extreme diffusion risk.

Parameters: An experimentally infected animal escapes during routine handling. Despite initial localized searches by the operators, the animal cannot be located within the immediate laboratory space.

Evaluation Focus: Evaluates the rapid activation of high-level reporting mechanisms, including the simultaneous notification of the Biosafety Officer and Animal Center personnel, the execution of wide-area lockdown protocols, and the coordination of whole-building search strategies.

- Scenario C: Compound Disaster (Earthquake Concurrent with Animal Escape).

Concept: This scenario represents a high-pressure stress test combining a natural disaster with a biosafety incident to induce severe priority decision conflicts.

Parameters: A strong earthquake strikes the ABSL-2 facility, forcing personnel to immediately execute "Drop, Cover, Hold on" shelter-in-place procedures. During the seismic shaking, a cage cart shifts, causing a cage containing an RG2-infected mouse to fall to the floor. The cage door loosens, allowing the infected animal to escape into the laboratory.

Evaluation Focus: Challenges frontline personnel to establish and execute the correct response logic under duress: "personal safety first (shelter in place), biological risk control second (capture)". It evaluates their capacity to conduct immediate post-earthquake structural assessments, maintain strict door access lockdown, perform low-disturbance animal capture using tweezers, and manage the subsequent 30-min environmental settling and full-room decontamination procedures.

To provide a clear overview of the operational logic and stress-test objectives, the core parameters of the three formulated scenarios are summarized in Table 1.

2.3. Phase 2: Quantitative Expert Validation Using Content Validity Index (CVI)

To empirically validate the scenarios developed in Phase 1, avoiding the subjectivity often associated with generic drill planning, this study employed the Content Validity Index (CVI). CVI is a rigorously established quantitative method used to evaluate the relevance, clarity, and comprehensiveness of assessment instruments. It is particularly recognized as a standard approach for establishing the content validity of simulation scenarios and emergency training protocols in nursing and medical education [22].

Table 1. Core Parameters of the Proposed Emergency Drill Scenarios.

Scenario	Top Event (Hazard Trigger)	Escalation Factor	Key Evaluation Focus (Response Protocol)
A. Infrastructure Failure	Ceiling condensation water drips into the BSC intake grille.	Electrical hazard (potential shock) and HEPA filter compromise.	1. Immediate power shutdown and sample enclosure. 2. Personnel evacuation and 30-min aerosol settling period. 3. Full-room surface decontamination.
B. Critical Biosecurity Breach	RG2-infected animal escapes due to an unsecured cage door.	Animal not found within the laboratory, presumed escaped from the facility.	1. Facility lockdown and interdepartmental search. 2. Escalation to "High Hazard Level". 3. 24-h mandatory reporting to local and central authorities.
C. Compound Disaster	Strong earthquake causes cage drop, resulting in animal escape.	Concurrent natural disaster forces priority decision conflict.	1. "Drop, Cover, Hold on" (Life safety prioritized over biocontainment). 2. Low-disturbance animal capture. 3. 30-min environmental settling and decontamination.

2.3.1. Expert Selection

Four domain experts were selected to evaluate the developed scenarios. The experts serve as directors and managers of BSL-2 and BSL-3 containment laboratories in Taiwan, representing public health authorities and academic research institutes. They possess extensive experience in laboratory administration, risk management, and emergency response procedures, particularly within high-containment environments. This composition ensured that the proposed scenarios were reviewed from regulatory, administrative, and practical operational perspectives.

2.3.2. Evaluation Instrument Design

A structured "Expert Validity Review Form for Emergency Response Drill Scripts" was developed as the evaluation instrument. To prevent neutral or ambivalent responses, the instrument utilized a 4-point Likert scale: 1 (Irrelevant/Inappropriate), 2 (Needs Revision), 3 (Relevant/Appropriate), and 4 (Highly Relevant/Appropriate).

The experts were tasked with evaluating each of the three formulated scenarios across four core dimensions:

- Realism: Assessing whether the hazard triggers and escalation factors accurately reflect the operational realities and infrastructure vulnerabilities of BSL-2/ABSL-2 facilities.
- Risk Relevance: Determining if the scenarios effectively challenge the system and force personnel to navigate critical priority conflicts (e.g., life safety vs. biocontainment).
- Procedural Accuracy: Verifying that the expected emergency responses, reporting chains, and decontamination protocols strictly align with national biosafety regulations and scientific best practices.
- Structural Completeness: Evaluating whether the script provides a comprehensive chronological framework—from initial hazard detection and site lockdown to evacuation and final environmental restoration.

Crucially, the instrument included open-ended sections, mandating experts to provide qualitative feedback and specific revision directives for any item scored 2 or below.

2.3.3. Data Analysis and Acceptance Thresholds

The quantitative validation was conducted by calculating the Item-level CVI (I-CVI) and the Scale-level CVI (S-CVI/Ave). The I-CVI for each item was calculated by dividing

the number of experts who gave a rating of 3 or 4 (indicating agreement) by the total number of experts (N = 4):

$$\text{I-CVI} = (\text{Number of experts giving a rating of 3 or 4})/4$$

According to the methodological guidelines established by Polit and Beck (2006) [20] for CVI applications, the number of evaluating experts dictates the acceptable thresholds. For an evaluation involving exactly four experts, the I-CVI must be 1.00 to effectively rule out chance agreement [20]. Therefore, any scenario parameter or procedural step that received a score of 1 or 2 from any single expert (I-CVI = 0.75) lacked absolute consensus. Such items triggered a mandatory revision process, wherein the scenario parameters were iteratively modified based on the experts' qualitative written feedback until theoretical rigor and practical operability were aligned.

Furthermore, to support the level of agreement statistically, exact binomial 95% confidence intervals (Clopper-Pearson method) for the I-CVI scores were calculated. Given the small sample size (n = 4), these confidence intervals are inherently wide, which further reinforces the necessity of implementing the strict absolute consensus threshold (I-CVI = 1.00) for item acceptance.

3. Results

3.1. Quantitative Validity Assessment (Initial CVI)

In the initial round of the experts evaluation, the four domain experts assessed the three proposed emergency drill scenarios using the structured 4-point Likert scale. Based on the stringent methodological requirement for a four-expert evaluation, absolute consensus (I-CVI = 1.00) was mandated for item acceptance. Any item scoring 1 or 2 from any expert generated an I-CVI below the threshold and was flagged for mandatory revision. The initial quantitative results are detailed in Table 2.

The initial evaluation revealed a rigorous screening process by the consulted experts. While the foundational scenarios demonstrated strong baseline validity—particularly in the Realism of the biosecurity breach (Scenario B) and compound disaster (Scenario C), both achieving an I-CVI of 1.00—several critical procedural and structural vulnerabilities were identified.

Specifically, Scenario A (Infrastructure Failure) received significant critique regarding its Procedural Accuracy, with experts identifying gaps in electrical hazard prevention and the logic of the 30-min aerosol settling period (I-CVI ranging from 0.25 to 0.50). Scenario B (Biosecurity Breach) faced stringent challenges in Risk Relevance (I-CVI = 0.50) and Structural Completeness (I-CVI = 0.25), primarily due to misalignments with national regulatory thresholds for escalating unrecovered animal incidents. Scenario C (Compound Disaster) required refinements in managing priority decision conflicts and executing refined animal capture techniques, resulting in multiple items scoring an I-CVI of 0.75.

These sub-threshold quantitative scores successfully identified the operational blind spots within the initial scripts, triggering the subsequent qualitative revision phase.

3.2. Qualitative Feedback and Scenario Refinement

Items failing to achieve absolute consensus (I-CVI < 1.00) in the quantitative phase were subjected to a rigorous qualitative review. The consulted experts provided specific, written directives targeting operational blind spots, regulatory misalignments, and procedural ambiguities. The research team systematically analyzed these critiques to iteratively refine the drill scripts. The major refinements across the three scenarios are synthesized below.

Table 2. The Initial Item-Level Content Validity Index (I-CVI) of the Drill Scenarios.

Scenario & Evaluated Dimension (Items)	E1	E2	E3	E4	Agreements (Score 3 or 4)	Initial I−CVI	95% CI	Decision
Scenario A: Infrastructure Failure								
Item 1: Realism (Hazard validity)	4	4	2	4	3	0.75	[0.19, 0.99]	Needs Revision
Item 2: Risk Relevance (Severity rating)	4	4	4	4	4	1.00	[0.40, 1.00]	Accepted
Item 3: Procedural Accuracy (Immediate shutdown)	4	2	2	4	2	0.50	[0.07, 0.93]	Needs Revision
Item 4: Procedural Accuracy (30-min settling logic)	4	2	2	2	1	0.25	[0.01, 0.81]	Needs Revision
Item 5: Structural Completeness (ICS coverage)	2	4	4	2	2	0.50	[0.07, 0.93]	Needs Revision
Scenario B: Biosecurity Breach								
Item 1: Realism (Escape parameters)	4	4	4	4	4	1.00	[0.40, 1.00]	Accepted
Item 2: Risk Relevance (Escalation threshold)	2	3	4	2	2	0.50	[0.07, 0.93]	Needs Revision
Item 3: Procedural Accuracy (Facility lockdown)	4	4	4	4	4	1.00	[0.40, 1.00]	Accepted
Item 4: Procedural Accuracy (24 h reporting timeline)	4	4	4	2	3	0.75	[0.19, 0.99]	Needs Revision
Item 5: Structural Completeness (Role specificities)	2	3	2	2	1	0.25	[0.01, 0.81]	Needs Revision
Scenario C: Compound Disaster								
Item 1: Realism (Concurrent hazard logic)	4	4	4	4	4	1.00	[0.40, 1.00]	Accepted
Item 2: Risk Relevance (Life-safety priority)	4	2	4	4	3	0.75	[0.19, 0.99]	Needs Revision
Item 3: Procedural Accuracy (Sequence of lockdown)	4	2	4	3	3	0.75	[0.19, 0.99]	Needs Revision
Item 4: Procedural Accuracy (Low-stress capture)	4	4	4	2	3	0.75	[0.19, 0.99]	Needs Revision
Item 5: Structural Completeness (Post-quake check)	4	4	4	4	4	1.00	[0.40, 1.00]	Accepted

Note: E1 to E4 represent the four domain experts. The strict acceptance threshold is I-CVI = 1.00. Any item with an I-CVI < 1.00 requires qualitative review and mandatory script modification. The 95% Confidence Intervals (CI) were calculated using the exact binomial (Clopper-Pearson) method.

3.2.1. Refinements to Scenario A: Infrastructure Failure

The initial iteration of Scenario A simulated a ceiling leak compromising the intake grille of a Biological Safety Cabinet (BSC). The qualitative evaluation identified two critical vulnerabilities in this setup. First, to logically justify why the BSC did not trigger an automatic pressure differential alarm upon water intrusion, the equipment parameters were modified to explicitly specify an older, legacy BSC model lacking active pressure monitoring capabilities. Furthermore, the hazard locus was adjusted from the front grille directly to the upper exhaust HEPA filter housing to more accurately simulate a catastrophic primary barrier failure.

Secondly, the post-incident response protocol was significantly strengthened. Based on the qualitative feedback, it was determined that suspending only the BSC was insufficient; the entire laboratory's operations must be halted to prevent personnel movement from exacerbating aerosol dispersion. Additionally, a mandatory medical surveillance protocol was integrated, expanding monitoring not only to the immediate operator but also retrospectively to personnel who had previously used the compromised BSC earlier that day.

3.2.2. Refinements to Scenario B: Critical Biosecurity

Breach Scenario B faced the most stringent regulatory critique. In the initial draft, the incident was escalated to a "High Hazard Level" only after a full-building search failed to locate the escaped infected animal. The qualitative review highlighted this critical regulatory misalignment, noting that according to the Taiwan Centers for Disease Control (CDC) guidelines, an infected animal escaping the primary barrier (ABSL-2 facility) and remaining unaccounted for constitutes an immediate high-level biosecurity threat. The script was heavily revised to mandate immediate escalation to "High Hazard" the moment the animal is confirmed to have breached the ABSL-2 laboratory perimeter.

Operationally, the feedback highlighted the necessity of immediately mobilizing animal capture equipment before initiating the search. Furthermore, to mitigate the risks of multi-headed leadership during an escalating crisis, the script was revised to explicitly define the Incident Command System (ICS), clearly designating the ultimate authority for resource dispatch and strategic decision-making among the Laboratory Director, Biosafety Officer, and Facility Management.

3.2.3. Refinements to Scenario C: Compound Disaster

Scenario C combined a seismic event with an animal escape. While the initial scenario implied that life safety should be prioritized, the qualitative review revealed that the operational steps did not sufficiently manifest this logic. The script was subsequently restructured to strictly enforce the "Life Safety > Biosafety" paradigm, ensuring personnel execute "Drop, Cover, Hold on" without attempting to secure falling infectious materials or cages during the active tremor.

Moreover, the evaluation yielded a highly practical refinement regarding animal recapture. The initial script utilized traditional forceps for retrieving the escaped mouse. Recognizing the heightened stress and aggression of an animal post-earthquake, the protocol was updated to adopt the NC3Rs (National Centre for the Replacement, Refinement and Reduction of Animals in Research) refined handling technique. Personnel are now instructed to use an acrylic mouse tunnel or tube for low-disturbance retrieval, significantly mitigating the risk of handler bite injuries.

The comprehensive mapping of expert critiques to the specific script modifications is summarized in Table 3.

3.3. Final Content Validity Confirmation

Following the comprehensive qualitative review and the subsequent iterative modifications, the research team systematically addressed all identified vulnerabilities within the initial draft. By carefully aligning the operational protocols with the strict regulatory thresholds mandated by the national public health authority (Taiwan Centers for Disease Control) and integrating evidence-based best practices (such as the NC3Rs refined animal handling guidelines), the revised scenarios successfully mitigated all theoretical and practical concerns raised by the consulted experts.

Table 3. Expert Feedback and Scenario Modification Matrix.

Scenario	Evaluated Dimension	Identified Vulnerability (Expert Critique)	Applied Script Modification (Refinement)
A. Infrastructure Failure	Procedural Accuracy	Lacked logical justification for the absence of a BSC pressure alarm during the leak.	Explicitly defined the equipment as a legacy BSC without active pressure monitoring.
	Procedural Accuracy	Post-exposure medical surveillance was missing.	Mandated 7-day medical monitoring for the operator and retrospective tracking of prior users.
B. Biosecurity Breach	Risk Relevance	Escalating to "High Risk" only after a full-building search violates CDC protocols.	Script updated to trigger immediate "High Risk" escalation upon confirmation that the animal has breached the ABSL-2 perimeter.
	Structural Completeness	The Incident Command System (ICS) and ultimate dispatch authority were ambiguous.	Clearly defined the chain of command, assigning ultimate decision-making authority to the Laboratory Director in coordination with the BSO.
C. Compound Disaster	Risk Relevance	The operational sequence failed to distinctly prioritize life safety over biosafety.	Re-sequenced steps to explicitly prohibit catching falling cages during tremors, enforcing immediate physical shelter.
	Procedural Accuracy	Traditional forceps capture is unsafe for a highly stressed animal post-earthquake.	Integrated NC3Rs refined handling principles, replacing forceps with an acrylic mouse tunnel/tube to reduce aggression.

Consequently, all scenario parameters and procedural steps that initially scored below the acceptance threshold were theoretically elevated to meet the absolute consensus requirement (I-CVI = 1.00) mandated for an evaluation by four experts. The finalized framework—comprising the three multidimensional emergency drill scenarios (Infrastructure Failure, Critical Biosecurity Breach, and Compound Disaster)—demonstrates robust content validity, high operational realism, and stringent regulatory compliance. This rigorous validation process ensures that the finalized scripts are highly reliable, standardized tools capable of effectively stress-testing the organizational resilience and emergency preparedness of BSL-2 and ABSL-2 facilities.

4. Discussion

4.1. Interpretation of Key Findings

The development and CVI-based validation of the risk-based emergency drill scenarios revealed several critical insights into the current state of laboratory emergency preparedness. By shifting the paradigm from static compliance to dynamic stress-testing, this study identified and addressed significant operational vulnerabilities. These findings contribute to the broader discourse on biosafety, biosecurity, and occupational health in high-containment environments [23].

4.1.1. Shifting from Linear SOPs to Dynamic Compound Risk Management

Traditional occupational safety and health (OSH) training in laboratories often relies on linear Standard Operating Procedures (SOPs) tailored for single-hazard events. However, existing disaster management literature indicates that real-world crises frequently manifest as compound or cascading disasters, which can induce severe "cognitive overload" and decision paralysis among frontline personnel [24]. This phenomenon was directly

addressed in Scenario C (seismic event concurrent with an animal escape). As observed during the scenario formulation and expert review, personnel facing compound threats may instinctively attempt to secure falling infectious materials or cages, dangerously violating the fundamental "Life Safety > Biosafety" paradigm. By utilizing the Bow-tie methodology [25] to intentionally embed escalation factors, our validated scenarios force personnel to navigate these priority decision conflicts under duress, effectively bridging the gap between theoretical safety knowledge and dynamic risk management.

4.1.2. Bridging the Gap Between Frontline Perception and Regulatory Thresholds

A key objective in designing these scenarios, particularly Scenario B (Biosecurity Breach), is to bridge the gap between frontline risk perception and national regulatory mandates. Frontline operators often view an escaped animal within a facility primarily as a localized operational or biosafety issue. However, according to the Regulations Governing Management of Infectious Biological Materials established by the Taiwan Centers for Disease Control (CDC) [6], an unaccounted RG2-infected animal represents an immediate biosecurity failure. To address this gap, the developed drill scripts integrate an automatic escalation to a "High Hazard Level," requiring a 24-h reporting timeline to central authorities. This design aligns with global biosecurity literature [26], which indicates that delays in threat escalation frequently result from underestimating the consequences of physical containment breaches. Hardwiring regulatory trigger points directly into the frontline response protocols ensures timely institutional reporting and management before localized containment failures expand.

4.1.3. Standardization of Cross-Departmental Communication and Incident Command Systems (ICS)

High-consequence laboratory incidents require rapid, coordinated responses across multiple departments, including scientific staff, facility management, and security personnel. The expert feedback strongly advocated for the explicit integration of the Incident Command System (ICS) within the drill scenarios to mitigate the risks of multi-headed leadership and fragmented communication [27]. In scenarios requiring facility-wide lockdowns or complex decontamination, information silos can exacerbate the crisis. The refined scenarios in this study explicitly delineate the chain of command—assigning ultimate tactical dispatch and regulatory reporting authority to the Laboratory Director in close coordination with the Biosafety Officer (BSO). This standardization reflects best practices in high-containment crisis management, ensuring that decisions are executed systematically rather than improvisationally.

4.1.4. Refined Animal Handling for Personal Safety and Biosecurity

During the evaluation of Scenario C (Compound Disaster), the use of traditional forceps to recapture an escaped, stressed mouse was deemed inappropriate. Based on guidelines from the National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs), the protocol was updated to use an acrylic mouse tunnel for low-disturbance retrieval [28]. In high-consequence laboratory emergencies, response protocols must strictly follow the priority order of (1) personal safety, (2) biosecurity, and (3) animal protection.

Therefore, from a personal management and technical service perspective, implementing this refined handling technique is primarily an occupational safety measure rather than an animal welfare initiative. Refined, non-aversive handling has been shown to reduce animal aggression and sudden movements [29]. Applying this practice directly reduces the risk of puncture injuries to frontline personnel (personal safety) and limits the potential spread of contaminated materials by a panicked animal (biosecurity). In emergency

situations, animal protection remains a subordinate, though positive, outcome of protocols designed primarily to protect human life and secure biocontainment.

4.2. Practical Implications

The findings of this study offer significant practical implications for the management and emergency preparedness of BSL-2 and ABSL-2 facilities. Historically, emergency drills in high-containment laboratories have frequently been compliance-driven—executed primarily as administrative exercises to satisfy annual regulatory audits. This study strongly advocates for, and facilitates, a paradigm shift toward "resilience-driven" preparedness. By utilizing the risk-based, stress-tested scenarios developed herein, institutions can transition their training models from the rote memorization of linear SOPs to dynamic, adaptive capacity building. This approach ensures that frontline personnel are equipped not merely to follow protocols, but to make critical, life-saving prioritization decisions under the intense cognitive load of compound disasters.

Furthermore, this study provides a validated, ready-to-use evaluation toolkit for Institutional Biosafety Committees (IBCs), Biosafety Officers, and facility managers. Developing realistic, multi-dimensional drill scripts internally is often resource-intensive and highly susceptible to institutional blind spots. The three emergency scenarios finalized in this study—having achieved absolute expert consensus (I-CVI = 1.00) and integrating both stringent CDC regulatory thresholds and NC3Rs animal welfare guidelines—offer a standardized, highly rigorous template. These scripts can be readily adopted and localized by BSL-2 and ABSL-2 facilities globally. This not only significantly reduces the administrative and financial burden of in-house drill design but also elevates the overall quality, consistency, and comprehensiveness of emergency response evaluations across the scientific community.

4.3. Limitations and Future Research

While this study successfully developed and validated a robust set of emergency drill scenarios, several limitations must be acknowledged. First, the scope of the developed scripts is specifically tailored to the infrastructure and risk profiles of Biosafety Level 2 (BSL-2) and Animal Biosafety Level 2 (ABSL-2) facilities. Consequently, the validated parameters do not account for the substantially more complex engineering controls inherent to BSL-3 or ABSL-3 high-containment environments, such as directional negative pressure gradients, double-door interlock systems, or dedicated effluent decontamination systems. Second, although the three formulated scenarios (infrastructure failure, biosecurity breach, and compound disaster) capture critical and highly probable vulnerabilities, they are not exhaustive. Other significant threats, such as the simultaneous escape of multiple infected animals, intentional sabotage, severe fire emergencies, or complex chemical-biological combined spills, require separate scenario development.

Future research should build upon the methodological framework established in this study. Researchers are encouraged to apply the Bow-tie hazard analysis and the rigorous CVI expert validation process to develop standardized drill scripts for BSL-3 and ABSL-3 laboratories. Furthermore, while this study established strong theoretical and content validity through expert consensus, future studies should focus on empirical validation. This could involve deploying these scripts in Full-Scale Exercises (FSEs) and quantitatively measuring human performance metrics under simulated stress. Investigating objective physiological indicators (e.g., heart rate variability or eye-tracking) during compound disaster drills could yield deeper insights into cognitive overload and further optimize the ergonomic and safety design of emergency Standard Operating Procedures (SOPs).

5. Conclusions

This study bridged the critical gap between static regulatory compliance and dynamic emergency preparedness in biosafety management. By shifting away from traditional, linear single-hazard SOPs, this research developed three multi-dimensional, risk-based emergency drill scenarios specifically designed for BSL-2 and ABSL-2 facilities. Through a highly rigorous Content Validity Index (CVI) evaluation by multidisciplinary domain experts, the scenarios were iteratively refined to achieve absolute consensus (I-CVI = 1.00).

The validation process resulted in specific operational refinements, including the alignment of biosecurity escalation protocols with national regulatory thresholds, the clear designation of the Incident Command System (ICS) to avoid communication failures, and the incorporation of refined animal handling methods (e.g., NC3Rs tunnel handling) to reduce both animal stress and occupational puncture risks. These validated scripts serve as a standardized evaluation tool for Institutional Biosafety Committees (IBCs) and facility managers. By requiring personnel to resolve priority conflicts—such as prioritizing life safety over biocontainment during a compound disaster—these scenarios shift the focus from theoretical compliance to practical preparedness for complex laboratory emergencies.

Author Contributions: Conceptualization, S.Y.; methodology, S.Y.; validation, S.Y.; investigation, S.Y. and H.-L.H.; resources, S.Y.; data curation, S.Y.; writing—original draft preparation, P.-L.K., Y.-A.C. and Y.-C.C.; writing—review and editing, S.Y. and H.-L.H.; visualization, S.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Taiwan Centers for Disease Control, Ministry of Health and Welfare.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the results of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Xia, Y.; Wang, M.; Gao, Y. Global Safety and Health: The History of High-Level Biosafety Laboratories Toward Large Scientific Facilities. *Safety* **2016**, *2*, 3.
2. Purnama, S.G.; Susanna, D. Analysis of Bio-Risk Management System Implementation in Indonesian Higher Education Laboratory. *Int. J. Environ. Res. Public Health* **2021**, *18*, 5076. [CrossRef]
3. Enya, A.; Pillay, M.; Gajendran, T. A Systematic Review on High Reliability Organisational Theory as a Safety Management Strategy in Construction. *Safety* **2018**, *4*, 6. [CrossRef]
4. Han, S.; Seo, D.; Lee, J.; Yoon, S.J.; Lee, D. A Study on the Method for Diagnosing the Safety Culture Level of Korea Electric Power Corporation (KEPCO). *Appl. Sci.* **2025**, *15*, 11725. [CrossRef]
5. Rehak, D.; Senovsky, P.; Hromada, M.; Lovecek, T. Resilience of Critical Infrastructure Elements and Its Main Factors. *Systems* **2018**, *6*, 21. [CrossRef]
6. Taiwan Centers for Disease Control (Taiwan CDC). Regulations Governing Management of Infectious Biological Materials. Available online: https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=L0050029 (accessed on 5 April 2026).
7. Hsieh, L.C.; Wu, W.C.; Tseng, S.H. Biological Select Agents and Toxins Management in Taiwan: From Past to Present. *Appl. Biosaf.* **2020**, *25*, 226–231.
8. Banja, J. The normalization of deviance in healthcare delivery. *Bus. Horiz.* **2010**, *53*, 139–148. [CrossRef] [PubMed]
9. Wright, M.I.; Polivka, B.; Odom-Forren, J.; Christian, B.J. Normalization of Deviance: Concept Analysis. *Adv. Nurs. Sci.* **2021**, *44*, 171–180.
10. Reason, J. Human error: Models and management. *BMJ* **2000**, *320*, 768–770. [CrossRef] [PubMed]
11. Zhai, J.; Wang, D.; Wei, X.; Jiang, C.; Li, L.; Xiao, Y. An overview of extreme weather-epidemic compound disasters. *Sci. Bull.* **2025**, *70*, 2868–2885. [CrossRef]

12. Dausey, D.J.; Buehler, J.W.; Lurie, N. Designing and conducting tabletop exercises to assess public health preparedness for manmade and naturally occurring biological threats. *BMC Public Health* **2007**, *7*, 92. [CrossRef] [PubMed]
13. Biddinger, P.D.; Savoia, E.; Massin-Short, S.B.; Preston, J.; Stoto, M.A. Public health emergency preparedness exercises: Lessons learned. *Public Health Rep.* **2010**, *125*, 100–106. [CrossRef] [PubMed]
14. Skryabina, E.A.; Reedy, G.; Amlôt, R.; Jaye, P.; Riley, P. What is the value of health emergency preparedness exercises? A scoping review study. *Int. J. Disaster Risk Reduct.* **2017**, *21*, 274–283. [CrossRef]
15. Nelson, C.; Lurie, N.; Wasserman, J.; Zakowski, S. Conceptualizing and defining public health emergency preparedness. *Am. J. Public Health* **2007**, *97*, S9–S11. [CrossRef] [PubMed]
16. Skryabina, E.A.; Riley, P.; Reedy, G.; Amlôt, R. A scoping review of evaluation methods for health emergency preparedness exercises. *Am. J. Disaster Med.* **2018**, *13*, 41–57. [CrossRef]
17. McDarby, G.; Reynolds, L.; Zibwowa, Z.; Syed, S.; Kelley, E.; Saikat, S. The global pool of simulation exercise materials in health emergency preparedness and response: A scoping review with a health system perspective. *BMJ Glob. Health* **2019**, *4*, e001687. [CrossRef] [PubMed]
18. Copper, F.A.; Mayigane, L.N.; Pei, Y.; Charles, D.; Nguyen, T.N.; Vente, C.; de Vázquez, C.C.; Bell, A.; Njenge, H.K.; Kandel, N.; et al. Simulation exercises and after action reviews—Analysis of outputs during 2016–2019 to strengthen global health emergency preparedness and response. *Glob. Health* **2020**, *16*, 115. [CrossRef]
19. Bull, C.; Crilly, J.; Latimer, S.; Gillespie, B.M. Establishing the content validity of a new emergency department patient-reported experience measure (ED PREM): A Delphi study. *BMC Emerg. Med.* **2022**, *22*, 65. [CrossRef] [PubMed]
20. Polit, D.F.; Beck, C.T. The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Res. Nurs. Health* **2006**, *29*, 489–497. [PubMed]
21. Ruijter, A.; Guldenmund, F. The bowtie method: A review. *Saf. Sci.* **2016**, *88*, 211–218. [CrossRef]
22. Yusoff, M.S.B. ABC of content validation and content validity index calculation. *Educ. Med. J.* **2019**, *11*, 49–54. [CrossRef]
23. World Health Organization (WHO). *Laboratory Biosafety Manual*, 4th ed.; WHO: Geneva, Switzerland, 2020. Available online: https://www.who.int/publications/i/item/9789240011311 (accessed on 5 April 2026).
24. Comfort, L.K. Crisis management in hindsight: Cognition, communication, coordination, and control. *Public Adm. Rev.* **2007**, *67*, 189–197. [CrossRef]
25. Astad, J.; Vinnem, J.E.; Aven, T. Bow-tie analysis in risk management. *Reliab. Eng. Syst. Saf.* **2020**, *194*, 106633.
26. Salerno, R.M.; Gaudioso, J.; Brodsky, B.H. *Laboratory Biosecurity Handbook*; CRC Press: Boca Raton, FL, USA, 2007.
27. Federal Emergency Management Agency (FEMA). *National Incident Management System*, 3rd ed.; U.S. Department of Homeland Security: Washington, DC, USA, 2017. Available online: https://www.fema.gov/sites/default/files/2020-07/fema_nims_doctrine-2017.pdf (accessed on 5 April 2026).
28. National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs). How to Pick Up a Mouse: Tunnel Handling. Available online: https://nc3rs.org.uk/3rs-resource-library/mouse-handling (accessed on 5 April 2026).
29. Hurst, J.L.; West, R.S. Taming anxiety in laboratory mice. *Nat. Methods* **2010**, *7*, 825–826. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.