Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Protective adequacy and institutional handoffs in urban heat governance

Violette Eljure Eljure

veljure@gmail.com

Independent Researcher

Article

Keywords: extreme heat, climate adaptation, heat-health governance, protective adequacy, policy implementation, institutional handoffs, artificial intelligence

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11297118/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Extreme heat governance is increasingly capable of detecting risk, but evidence that protection changes when risk changes remains scarce. Here I develop Risk-to-Protection Translation, a six-state framework separating risk change, institutional recognition, authorized decision, operational delivery, population reach and protection. Comparative process tracing of Ahmedabad and Delhi, with Barcelona as a mechanism probe, shows that protective functions move further downstream when responsibility is clear and the responsible actor controls the required action. Translation becomes harder to verify as functions depend on organizational handoffs or third-party compliance without compensating coordination, monitoring or enforcement. The hardest downstream gaps appear in pathways serving populations with limited private alternatives or limited control over exposure. These findings distinguish sustained implementation from sustained protective adequacy. I derive a prospective AI-enabled Risk-to-Protection Standard that verifies downstream protective states and flags unresolved gaps while preserving human accountability. Prediction is therefore only the first stage of climate protection.

Introduction

Extreme heat is becoming a governance test as much as a forecasting problem. Climate change is increasing heat-related risks, while urban exposure, night-time heat, housing conditions, occupation, electricity reliability and access to private cooling shape who can act on an alert and who cannot.[1] Heat-health guidance has responded by expanding beyond warning systems to include governance, at-risk populations, communication, health-system resilience, exposure reduction, surveillance, monitoring and learning.[2] Recent international assessment of Heat Action Plans likewise shows that heat governance is a distinct and still underemphasized dimension of urban action.[3]

Yet the policy literature still has a downstream measurement problem. A warning can be accurate without changing a worker's schedule. A cooling facility can exist without being open when exposure peaks. Water can be distributed while administrative counts do not establish how many distinct people were reached. A ministry can issue an employer advisory while actual protection depends on thousands of worksites and an inspection system. In each case, the upstream action is real, but the protective condition experienced by the exposed population remains uncertain.

This problem is not solved by asking whether a plan exists or whether an intervention has been maintained. Urban climate research is increasingly concerned with delivery, including the configurations through which climate action moves from ambition into implementation.[4] Health adaptation research similarly emphasizes pathways, timing, barriers and delays as climate risks intensify.[5] These advances create a sharper question: once a material change in risk is recognized, what institutional transitions must occur before protection itself changes?

Here I address that question through the concept of protective adequacy: the correspondence between the current level and distribution of risk and the protection that is actually available to the population an

intervention is intended to protect. Protective adequacy is deliberately downstream. It cannot be inferred from policy adoption, implementation fidelity, service availability or institutional persistence alone. I operationalize the passage from updated risk to updated protection as Risk-to-Protection Translation (RPT), a six-state sequence that separates risk change, institutional recognition, authorized decision, operational delivery, population reach and evidence of protection.

The paper makes three linked claims. First, translation is more likely to advance when responsibility for the next protective step is clear and the responsible actor has authority and operational control over the required action. Second, organizational and third-party handoffs do not cause failure by themselves, but they create transitions that require coordination, monitoring, incentives or enforcement if downstream states are to remain observable and actionable. Third, the consequences of equivalent delay are greatest where exposed populations have few private alternatives and little control over the conditions needed to reduce exposure.

The empirical design uses comparative process tracing. Ahmedabad is a positive endpoint case because unusually strong evidence links local heat-risk knowledge, institutional action and a change in the temperature-mortality relationship. Delhi provides the principal causal leverage through within-government comparison across warning, cooling, water, health preparedness and occupational protection. Barcelona is used as a focused mechanism probe showing why physical availability and operational availability are not equivalent. The analysis is not a ranking of cities. It asks which transition had to occur next, who controlled it, whether it occurred, and what evidence supports that inference.

The argument also has a forward-looking implication. Artificial intelligence is rapidly improving heat prediction, vulnerability mapping and climate decision support, but recent reviews show that operational implementation, response and recovery remain much less developed than prediction and risk identification.[6,7] I therefore derive an AI-enabled Risk-to-Protection Standard as a prospective institutional design: AI can help reconcile data, detect missing downstream states and flag unresolved gaps, but it cannot substitute for public authority, resources, enforcement or rights. The empirical contribution remains the institutional mechanism; the AI architecture is a testable implication rather than a validated intervention.

Results

Protective adequacy requires evidence beyond implementation

The first result is conceptual but empirically consequential: the protective endpoint cannot be safely inferred from upstream administrative states. Institutional-fit and adaptation-barrier scholarship already establishes that institutions can become misaligned with changing environmental problems and that barriers can arise throughout adaptation processes.[8–11] Policy-capacity and dynamic-capability scholarship explains the competences and organizational routines that enable governments to learn and

reconfigure.[12,13] Implementation science likewise distinguishes adoption, reach, effectiveness, implementation and maintenance, and explicitly recognizes dynamic context and adaptation over time.[14–16] None of these literatures justifies treating an implemented or sustained intervention as evidence that protection remains sufficient when the hazard itself is moving.

Outcome evidence is beginning to strengthen. A 2025 multi-country European analysis associated implementation of heat prevention plans with a 25.2% reduction in excess deaths attributable to extreme heat, while a recent systematic review of heat-health warning systems emphasizes the need to connect warning design more closely to health outcomes and local vulnerability.[17,18]

RPT therefore separates six observable states (Fig. 1). T0 is a material change in risk or protective requirement. T1 is institutional recognition of that change. T2 is an authorized decision that changes what should happen. T3 is a verified operational change in delivery. T4 is usable access by or receipt among the intended population. T5 is evidence that exposure or heat-related health risk changed in the intended direction. The rule is conservative: no state is inferred from the state before it. When downstream evidence is missing, the status remains unknown rather than being coded as either failure or success.

This distinction yields two diagnostic constructs for future measurement. Risk-to-Protection Lag is the elapsed time between a verified change in protective requirement and a verified downstream protective state. The Protection Gap is the difference between the protective condition required or formally authorized and the condition that can be verified in delivery or reach. These are not validated indices in the present study. Their purpose is to identify where a protective chain becomes slow or incomplete and what evidence would be needed to test that claim prospectively.

Comparative evidence supports a handoff mechanism

Across the cases, the most consistent pattern is not that distributed governance fails, but that longer protective pathways require compensating institutional controls. Classic implementation research anticipated this problem by showing how action becomes more difficult as decision points accumulate, while climate-adaptation research repeatedly identifies fragmented mandates and coordination problems.[19,20] RPT narrows that general insight to a transition-level proposition: each transfer raises questions of responsibility, authority, observability and correction.

Figure 2 distinguishes a short pathway, in which a public actor can translate risk information relatively directly into action, from a longer pathway that crosses agencies, facilities, operators or employers before reaching the exposed population. Longer pathways are not intrinsically inferior. They become fragile when no actor can verify whether the next required state occurred or compel correction when it did not.

The cases provide graded support for this mechanism. Ahmedabad demonstrates that a pathway can connect local risk evidence, formal planning, coordinated implementation and a measurable health endpoint. Delhi shows substantial formal recognition of heat risk but more difficult downstream verification in functions that depend on facilities, distributed operations, employers or inspection. Barcelona shows that even a large infrastructure network requires adjustment of operating schedules before physical availability becomes usable access. A recent Western Cape study provides an external plausibility check: role ambiguity, fragmented responsibility and weak coordination were identified as barriers to cross-sector management of heat-related health risk in a different institutional setting.[21]

Table 1
Asymmetric case design. Cases provide different forms of causal leverage and are not used to rank cities.

Case	Inferential role	Primary leverage
Ahmedabad	Positive endpoint case	Local risk evidence, coordinated implementation and peer-reviewed post-plan mortality evaluation
Delhi	Principal mechanism case	Within-government comparison of functions with different handoff and operational-control structures
Barcelona	Focused mechanism probe	Difference between physical shelter availability and operating availability during high-exposure periods
Western Cape	External plausibility check	Independent evidence on role clarity and cross-sector coordination outside the core cases

Ahmedabad reaches an observable protection endpoint

Ahmedabad's Heat Action Plan was developed after the severe 2010 heatwave using local temperature-mortality evidence, hospital and ambulance information, surveys of exposed populations, stakeholder engagement and probabilistic forecasting.[22] The plan linked warning thresholds to public communication, health-system preparation and interagency roles, creating an observable chain from risk knowledge to organized action.

The case is especially valuable because the downstream endpoint was subsequently evaluated. Hess and colleagues compared temperature-mortality relationships before and after implementation and estimated 1,190 average annualized avoided deaths in the post-plan period, with a 95% confidence interval of 162 to 2,218.[23] The evaluation is observational and does not isolate the causal effect of each component, but it provides stronger T5 evidence than is available for the other cases. It supports the inference that an institutional pathway can travel far enough for a change in protection to become empirically visible.

Ahmedabad does not prove that role clarity or coordination alone caused the mortality reduction. Leadership, resources, public attention and concurrent changes remain plausible contributors. Its role in

the comparison is narrower: it demonstrates that risk recognition, decision, delivery and a measurable health endpoint can be connected without treating every upstream activity as equivalent to protection.

Delhi reveals within-government variation

Delhi provides the main mechanism test because multiple protective functions operate within the same broader administrative environment. Heat Wave Action Plans document forecast-linked recognition, vulnerable-population attention and assigned responsibilities across warning, health preparedness, water, cooling and occupational protection.[24,25] This common formal architecture makes within-city variation more informative than a simple comparison of cities with and without plans.

Warning and communication follow relatively short administrative pathways. Once a forecast crosses an alert threshold, authorities can issue public messages or formal instructions without requiring a large number of external actors to change their behavior. The evidence reviewed therefore supports strong T1-T2 architecture for these functions, although population-level receipt is not systematically measured here.

Cooling access requires more transitions. A facility may be listed in a plan, but usable protection depends on location, opening hours, staffing, water, accessibility and public awareness. A targeted 2026 audit of seven cooling centres across seven Delhi revenue districts documented the practical difficulty of locating some centres, variation in hours and accessibility, and very low worker awareness in its non-representative sample.[26,27] The audit interviewed 70 informal workers at 17 sites and combined field visits with right-to-information requests. It cannot estimate citywide prevalence, but it can identify downstream states that were difficult to verify and the institutional handoffs associated with them.

Drinking-water provision shows a different measurement problem. Administrative records can document quantities distributed without establishing how many distinct people actually received usable protection. The 2026 audit reported examples in which counts of distributed objects were close to reported beneficiary totals, illustrating why outputs and unique population reach should not be treated as interchangeable.[27] This is not proof of systematic mismeasurement across Delhi; it is direct evidence that a reach metric can become ambiguous at the point where delivery is translated into beneficiary counts.

Occupational heat produces the longest and most authority-fragmented pathway in the comparison. National and local advisories can specify measures such as altered work hours, water, shade and rest, but the immediate conditions of exposure are controlled by employers, contractors and site supervisors. In April 2026, India's Ministry of Labour and Employment issued a nationwide heatwave advisory urging states and employers to implement protective measures.[28] Yet the Delhi field audit reported that most interviewed workers had received no state-provided protective equipment and that an information request to the Chief Inspector of Factories returned no heat-specific inspection drive for the period examined.[27] Again, these observations are not representative estimates. Their analytical value lies in the

location of the gap: formal guidance exists upstream, while realized protection depends on a separate enforcement and employer-control pathway.

This pattern is consistent with evidence from the Centre for Science and Environment, which identifies outdoor workers and other highly exposed groups as requiring targeted protection in Delhi, including cooling breaks and staggered working hours.[29] It is also consistent with the institutional contrast provided by California's enforceable outdoor-heat standard, which specifies water, shade, preventive cool-down rest and high-heat procedures.[30] The comparison does not imply that one jurisdiction can be transplanted into another. It shows that occupational protection requires an enforceable pathway beyond risk communication.

Table 2
Delhi transition matrix. Unknown is an evidentiary status, not a coded failure.

Function	Pathway	Downstream evidence in reviewed record	Interpretive status
Warning / alerting	Short to moderate	Forecast-linked recognition and formal activation architecture are documented; systematic population receipt is not measured here	Strongest evidence at T1-T2
Cooling access	Moderate	Facilities specified; targeted 2026 audit documents uneven locatability, hours and awareness in audited sites	T3-T4 partial; T5 unknown
Drinking water	Moderate	Some delivery documented; beneficiary counting can conflate outputs with distinct people reached	T3 observed in some sites; T4 uncertain
Health preparedness	Moderate	Protocols and readiness responsibilities specified; event-level verification is incomplete in reviewed evidence	T2 strong; downstream states partly unknown
Occupational protection	Long / high handoff	Protective measures formally specified; targeted audit reports weak worker-level provision and no heat-specific inspection drive in responding record	High enforcement dependence; T5 unknown

Barcelona separates availability from access

Barcelona provides a narrower test of the distinction between infrastructure and usable access. In 2025 the city reported a network of nearly 400 climate shelters, with broad geographic coverage. It also increased the number operating in August to 256 on weekdays and 189 on weekends and kept additional libraries open during the holiday period.[31] The adjustment is not evidence of failure. It is evidence that a large physical network still has a temporal operating dimension that must be recalibrated when heat exposure and service availability diverge.

The case therefore supports the distinction between T3 and T4. A cooling space can physically exist while not being usable at the time a person needs it. Barcelona's reinforcement of summer schedules

illustrates a feedback loop in which an access condition is observed and the operating arrangement is changed. That is precisely the kind of downstream adjustment that an implementation count alone would miss.

Protection dependence sharpens equity implications

The same institutional delay does not have the same consequence for everyone. The comparison supports a distributional concept I call protection dependence: risk is especially consequential when high exposure coincides with limited private substitution and limited control over the action needed to reduce exposure.

Outdoor workers are the clearest example. A worker may receive and understand a warning but still lack authority to change hours, pace, shade, rest or whether work stops. An older adult may know that heat is dangerous but lack air conditioning or transport to a cooling facility. A person experiencing homelessness may receive an alert without having a private space in which to act on it. In these settings, communication can succeed while protection fails to become controllable by the person at risk.

This shifts the equity question from targeting alone to institutional dependence. Targeting identifies who should be prioritized. Protection dependence identifies where upstream information is least likely to be sufficient because the protective condition is controlled by another actor or institution. In the cases reviewed, the most consequential downstream gaps are therefore not simply those with the longest administrative chains, but those in which exposed populations have the least ability to compensate privately when the chain stalls.

AI exposes a forecast-to-protection gap

The rapid growth of artificial intelligence in heat governance makes the downstream distinction more urgent. A 2026 systematic review of AI and machine-learning applications in heat-risk management found a strong concentration in prediction and prevention, with much thinner evidence for preparedness, response and recovery.[6] A second review of 102 publications on AI for urban extreme heat similarly found strong activity in forecasting, monitoring and risk identification but limited causal and multi-stage modeling across the full heat-risk chain.[7] New machine-learning systems are also improving localized early warning by integrating climate projections, social vulnerability and other environmental drivers.[32,33]

These advances are valuable, but they create a forecast-to-protection paradox. A government can become increasingly precise about where risk is highest while remaining unable to verify whether the protective action expected in response actually occurred. More accurate risk intelligence does not, by itself, create operating hours, staffing, water, inspections, employer compliance or access.

I therefore derive an AI-enabled Risk-to-Protection Standard as a prospective design implication (Fig. 3). The system would connect risk sensing to five downstream verification layers: decision state, delivery,

reach, protection and learning. A Protection Failure Alert would be generated when current risk crosses a predefined threshold and an expected downstream state is missing, late, internally inconsistent or below the required condition. The alert would identify the affected function, the evidence gap, the responsible actor, source freshness and the authority responsible for escalation.

The proposed architecture is intentionally asymmetric. AI may help reconcile multiple data streams, identify anomalies, prioritize verification and surface missing evidence. It should not infer protection merely because no failure was reported. If facility status is stale, the state should remain unknown. If employer compliance has not been observed, absence of complaints should not be converted into compliance. If administrative counts cannot distinguish objects distributed from unique people reached, the system should preserve that uncertainty.

This design follows established principles for responsible AI in health and public decision systems, including human accountability, transparency, validity, fairness, privacy and continuous monitoring.[34,35] Recent work on trustworthy AI for climate services likewise emphasizes human-in-the-loop design, traceability and explicit treatment of uncertainty.[36] These safeguards matter because the AI layer could otherwise become another handoff: a technical system that produces signals without clear ownership of the corrective action.

The RPS is not an evaluated intervention in this study. Its testable value is prospective. A future pilot should pre-register heat thresholds, expected protective actions and transition deadlines; compare conventional monitoring with RPS-assisted monitoring; and evaluate time to gap detection, time to corrective action, verified reach, unresolved alerts, subgroup false negatives and administrative burden. The central outcome is not predictive accuracy alone. It is whether downstream verification shortens Risk-to-Protection Lag or reduces a Protection Gap without transferring public accountability to an algorithm.

Discussion

The comparison changes what should count as evidence of climate adaptation. Adoption, persistence and service delivery remain important achievements, but they occur upstream of the outcome that ultimately matters: whether the protective condition experienced by exposed people remains adequate as risk changes. This endpoint discipline is particularly important for extreme heat because both hazard and vulnerability can move faster than institutional routines.

The findings complement, rather than compete with, recent work in npj Climate Action. The international assessment of Heat Action Plans identifies governance as an underdeveloped component of heat action.[3] The climate-delivery framework developed by Sudmant and colleagues identifies recurring modes through which urban interventions move from ambition to delivery.[4] Health-adaptation pathway research emphasizes timing, implementation delay and the possibility that interventions become insufficient under different development and warming trajectories.[5] RPT adds a narrower downstream

question to this emerging agenda: after a plan contains the right elements and a delivery mode is activated, what evidence is required before a change in protection can be claimed?

The cases suggest that the answer is function-specific. Aggregate assessments of city capacity or plan quality can obscure variation inside the same administration. Warning systems may travel through relatively short chains; cooling access depends on facilities and operating schedules; occupational protection depends on employers, rights, monitoring and enforcement. The relevant unit of governance is therefore not the city alone but the transition required by a particular protective function. This is why a city can look administratively capable and still have a weak protective pathway for a subgroup whose exposure is controlled elsewhere.

The handoff mechanism is deliberately bounded. The evidence does not establish that handoffs are the sole cause of uneven translation, nor that fewer actors are always preferable. Resources remain a serious explanation for water, facilities, staffing and inspection. Political attention can shape enforcement. Information quality matters at the risk-recognition stage and again when governments try to observe delivery and reach. The stronger inference is conditional: where the next required action sits with another organization or private actor, downstream translation becomes harder to verify unless the institutional design makes that transfer visible and governable.

The study also identifies a measurement asymmetry. Risk information is increasingly sophisticated, while evidence of delivery, reach and protection often becomes thinner as the chain moves downstream. That asymmetry should itself be treated as a governance problem. Monitoring is not merely an evaluation exercise after a heat season. In a fast-moving hazard, it is part of implementation because authorities need to know whether the state they authorized actually exists while there is still time to correct it.

This has direct implications for equity. Climate-health policy often identifies vulnerable groups correctly but assumes that communicating risk or offering a service is enough to make protection available. Protection dependence shows why this can fail. The highest-stakes pathways are those in which exposed people cannot substitute privately and cannot control the conditions required to reduce exposure. For these groups, the reliability of institutional handoffs is not an administrative detail; it is part of the protective intervention.

The artificial-intelligence implication follows from the same logic. AI can increase the speed and granularity of risk intelligence and may help governments integrate operational data that are currently fragmented. But the appropriate benchmark is not whether a dashboard is more sophisticated. It is whether the system makes downstream failure visible earlier, improves corrective action and does so without introducing new inequities or obscuring responsibility. In that sense, AI should be evaluated as an institutional intervention rather than as a prediction model alone.

The study has important limits. The evidence is asymmetric across cases: Ahmedabad has a peer-reviewed health-outcome evaluation, Delhi offers stronger within-government variation but limited

representative evidence of downstream states, and Barcelona is intentionally a narrow mechanism probe. The Delhi audit is targeted rather than citywide and is used only for process observations. The study does not estimate a causal treatment effect of handoff burden on health outcomes, and the proposed Risk-to-Protection Lag and Protection Gap have not been standardized or validated across cities. The AI-enabled RPS remains a prospective design, not a demonstrated system.

Those limitations define the next test. A multi-city prospective design should pre-register transition-specific expectations before the heat season, time-stamp movement from T0 through T5, record handoffs and authority, and compare functions with different pathway complexity while accounting for resources, political attention and baseline administrative capacity. An RPS pilot should then test whether downstream verification improves detection, correction and reach relative to existing monitoring, with explicit auditing for populations that are most likely to be missing from digital or administrative data.

Extreme heat governance is moving into a phase in which better prediction is no longer enough. The more precisely governments can identify danger, the harder it becomes to justify not knowing whether protection followed. Protective adequacy provides a downstream standard for that question: did the institution merely recognize changing risk, or did it change the conditions experienced by the people who depended on public action?

Methods

Study design

The study uses comparative process tracing to examine how recognized changes in heat risk move through institutional systems. It does not estimate a treatment effect of RPT on mortality or model the probability of transition statistically. The unit of analysis is a protective function at a specific institutional transition. Protective functions include warning and communication, cooling access, drinking water, health preparedness and occupational heat protection.

For each function, the protocol records six states: T0, a material change in risk or protective requirement; T1, institutional recognition of that change; T2, an authorized decision; T3, verified operational delivery; T4, usable access by or receipt among the intended population; and T5, evidence that exposure or heat-related health risk changed in the intended direction. The protocol also records the actor formally responsible for the transition, the actor with operational control over the required action, the evidence of expected and observed action, plausible binding constraints, rival explanations and a qualitative confidence assessment. The complete coding dictionary is reported in Supplementary Table 1.

Case selection

Case selection is asymmetric by design. Ahmedabad is included as a positive endpoint case because it provides unusually strong evidence linking local heat-risk information, institutional design and a post-plan mortality evaluation.[22,23] Delhi is the principal mechanism case because the same broader

governing environment contains protective functions with different levels of organizational and third-party dependence.[24–30] This within-government variation places more pressure on explanations based only on aggregate city capacity, wealth or formal political commitment. Barcelona is a focused mechanism probe of operational availability because its climate-shelter network documents both broad physical coverage and deliberate adjustment of opening schedules during high-exposure periods.[31] The Western Cape evidence is used only as an external plausibility check on role clarity and coordination and is not pooled as a co-equal case.[21]

Evidence and source hierarchy

The evidence base combines official Heat Action Plans and administrative documents, municipal and regulatory materials, peer-reviewed evaluations, public-health guidance, and targeted independent evidence for downstream states not documented in official sources. Primary government documents are used to establish formal responsibilities, thresholds and authorized actions. Peer-reviewed studies are prioritized for health outcomes and broader empirical claims. Independent field reporting is used only for process observations unavailable in official documents and is explicitly bounded by its sampling design.

The Delhi analysis follows this rule closely. The 2026 field audit covered seven revenue districts, audited seven cooling centres, interviewed 70 informal workers across 17 sites and used right-to-information requests to examine implementation.[26,27] It is not treated as representative of Delhi's population or all facilities. It is used to identify specific downstream states, measurement problems and handoffs that were directly observed. Where independent observations and formal documents diverge, the discrepancy is treated as a question for verification rather than as automatic proof that the formal system failed everywhere.

The empirical cutoff for the manuscript is 30 September 2026. Documents and studies published after this date are not included in case coding. The 2026 npj Climate Action literature published before the cutoff is used to position the contribution but does not alter case coding.

Coding rules and confidence

Each T-state is coded as observed, partial or unknown. Observed requires direct evidence that the state occurred for the function and period under review. Partial indicates credible evidence of some components without sufficient evidence for the full state. Unknown is used when the available record does not permit a defensible inference. Unknown is not coded as failure and is never converted into protection.

The protection endpoint is function-specific. For health interventions, T5 can be supported by mortality or morbidity evidence. For cooling, it would require evidence that the relevant exposure condition changed, not simply that a facility existed. For occupational heat, it would require verified control of workplace exposure. Attendance, contact, objects distributed and messages sent are treated as delivery or reach indicators unless evidence supports a stronger protective inference.

Confidence is graded as high, moderate or low. High confidence requires clear temporal ordering and multiple credible sources, preferably including primary documentation. Moderate confidence reflects credible but incomplete observation of a transition or mechanism. Low confidence is reserved for indirect or single-source inference. These confidence labels describe the evidentiary strength of a transition-level claim; they are not probabilities.

Mechanism test and rival explanations

The mechanism test evaluates whether responsibility, authority, operational control and handoff structure account for variation in downstream translation. The observable expectation is that, within the same governing environment, functions with shorter or more governable pathways should show more complete or faster movement through T1-T4 than functions requiring additional organizations, facilities, employers or enforcement. A longer pathway that translates successfully should reveal compensating coordination, monitoring, incentives or enforcement. Supplementary Table 2 summarizes the rival-explanation signatures used in the analysis.

Four rival or complementary explanations are considered. Aggregate administrative capacity predicts more uniform performance across functions within the same government. Resource scarcity predicts gaps where the next transition requires unavailable staff, infrastructure, supplies or inspection capacity. Political commitment predicts stronger movement where formal agenda attention is matched by implementation effort. Information quality predicts failure at T0-T1 when risk is poorly recognized and can also explain downstream gaps when delivery, access or compliance are not observable. The handoff explanation gains support only where the observed bottleneck appears at a transfer of responsibility and rival explanations do not provide a more complete account.

No statistical hypothesis tests are performed, and no missing observation is treated as a negative outcome. The objective is mechanism discrimination rather than effect-size estimation. Supplementary Table 3 reports the full Delhi transition evidence matrix used to discipline the within-city comparison.

Declarations

Ethics and responsible use of artificial intelligence

The study analyzes published and publicly available secondary materials. It does not involve original collection of identifiable human-participant data, patient-level data or animal data; ethics approval and participant consent are therefore not applicable. Where published field studies involved interviews, this manuscript relies only on their publicly reported aggregate findings and does not access underlying identifiable records.

During manuscript preparation, OpenAI ChatGPT was used to support literature discovery, conceptual organization, drafting support, structural revision and language refinement. The author independently reviewed the sources and references used in the manuscript, made all analytical and interpretive

decisions, revised the text, and assumes full responsibility for the submitted content. No AI system is listed as an author or used as an autonomous decision-maker in the study.

Data Availability

All materials analyzed in this study are publicly available sources identified in the reference list and Supplementary Information. No new patient-level, individual-level, proprietary or restricted dataset was generated. The transition coding reported in the article is reproduced in the Supplementary Information so that the evidentiary basis of each inference can be inspected and extended.

Code Availability

No custom computer code or algorithm was used to generate the empirical results reported in this study.

Competing Interests

The author declares no financial or non-financial competing interests.

Author Contribution

V.E.E. conceived the study; developed the conceptual framework, research question and methodology; curated and analysed the documentary evidence; interpreted the findings; prepared the figures; and drafted, critically revised and approved the manuscript. V.E.E. is solely responsible for the final content.

Acknowledgements

This study received no funding.

References

1. Intergovernmental Panel on Climate Change. Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report (Cambridge University Press, 2022).
2. World Health Organization Regional Office for Europe. Heat-Health Action Plans: Guidance. 2nd ed. (WHO Regional Office for Europe, 2026).
3. Sobhaninia, S. et al. An international assessment of heat action plans. npj Clim. Action (2026). https://doi.org/10.1038/s44168-026-00416-9.
4. Sudmant, A. et al. From urban climate ambition to delivery: a framework of climate delivery modes. npj Clim. Action (2026). https://doi.org/10.1038/s44168-026-00408-9.
5. Ebi, K. L. et al. From frameworks to futures: health adaptation pathways across the Shared Socioeconomic Pathways. npj Clim. Action (2026). https://doi.org/10.1038/s44168-026-00414-x.

6. Easton-Calabria, L. & Patel, K. Bridging the implementation gap: artificial intelligence and machine learning applications in heat risk management. Clim. Risk Manag. 53, 100827 (2026). https://doi.org/10.1016/j.crm.2026.100827.
7. Rui, J., Shabrina, Z. & Gong, W. Artificial intelligence applications in urban extreme heat management: a systematic review of forecasting, monitoring, mitigation and decision support. Environ. Impact Assess. Rev. 119, 108363 (2026). https://doi.org/10.1016/j.eiar.2026.108363.
8. Moser, S. C. & Ekstrom, J. A. A framework to diagnose barriers to climate change adaptation. Proc. Natl Acad. Sci. USA 107, 22026–22031 (2010). https://doi.org/10.1073/pnas.1007887107.
9. Biesbroek, G. R., Klostermann, J. E. M., Termeer, C. J. A. M. & Kabat, P. On the nature of barriers to climate change adaptation. Reg. Environ. Change 13, 1119–1129 (2013). https://doi.org/10.1007/s10113-013-0421-y.
10. Young, O. R. The Institutional Dimensions of Environmental Change: Fit, Interplay, and Scale (MIT Press, 2002).
11. Epstein, G. et al. Institutional fit and the sustainability of social-ecological systems. Curr. Opin. Environ. Sustain. 14, 34–40 (2015). https://doi.org/10.1016/j.cosust.2015.03.005.
12. Wu, X., Ramesh, M. & Howlett, M. Policy capacity: a conceptual framework for understanding policy competences and capabilities. Policy Soc. 34, 165–171 (2015). https://doi.org/10.1016/j.polsoc.2015.09.001.
13. Mazzucato, M. & Kattel, R. Market-Shaping States: A New Theory of Public Sector Capacities and Capabilities. UCL Institute for Innovation and Public Purpose Working Paper 2026-03 (2026).
14. Glasgow, R. E., Vogt, T. M. & Boles, S. M. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am. J. Public Health 89, 1322–1327 (1999). https://doi.org/10.2105/AJPH.89.9.1322.
15. Chambers, D. A., Glasgow, R. E. & Stange, K. C. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. Implement. Sci. 8, 117 (2013). https://doi.org/10.1186/1748-5908-8-117.
16. Shelton, R. C., Chambers, D. A. & Glasgow, R. E. An extension of RE-AIM to enhance sustainability: addressing dynamic context and promoting health equity over time. Front. Public Health 8, 134 (2020). https://doi.org/10.3389/fpubh.2020.00134.
17. Urban, A. et al. The effectiveness of heat prevention plans in reducing heat-related mortality across Europe. Environ. Res. Lett. 20, 124071 (2025). https://doi.org/10.1088/1748-9326/ae2775.
18. Chandra, N. S. V. S. & Lee, J. K. W. A systematic review of heat health warning systems: enhancing the framework towards effective health outcomes. Curr. Environ. Health Rep. 12, 31 (2025). https://doi.org/10.1007/s40572-025-00496-5.
19. Pressman, J. L. & Wildavsky, A. Implementation. 3rd ed. (University of California Press, 1984).
20. Runhaar, H. A. C., Wilk, B., Persson, A., Uittenbroek, C. J. & Wamsler, C. Mainstreaming climate adaptation: taking stock about what works from empirical research worldwide. Reg. Environ. Change 18, 1201–1210 (2018). https://doi.org/10.1007/s10113-017-1259-5.

21. Quintana, A. V., Gilson, L., Kovats, S., Wright, C. Y. & Mayhew, S. H. The implementation gap: cross-sector management of heat-related health risks in Western Cape, South Africa. PLOS Glob. Public Health 5, e0004699 (2025). https://doi.org/10.1371/journal.pgph.0004699.
22. Knowlton, K. et al. Development and implementation of South Asia's first heat-health action plan in Ahmedabad, Gujarat, India. Int. J. Environ. Res. Public Health 11, 3473–3492 (2014). https://doi.org/10.3390/ijerph110403473.
23. Hess, J. J. et al. Building resilience to climate change: pilot evaluation of the impact of India's first heat action plan on all-cause mortality. J. Environ. Public Health 2018, 7973519 (2018). https://doi.org/10.1155/2018/7973519.
24. Delhi Disaster Management Authority. Heat Wave Action Plan for Delhi 2024 (Government of NCT of Delhi, 2024).
25. Delhi Disaster Management Authority. Heat Wave Action Plan for Delhi 2025 (Government of NCT of Delhi, 2025).
26. The Leaflet. Auditing Delhi's Heat Action Plan: A Ground Study, Part I (9 August 2026). https://theleaflet.in/governance-and-policy/auditing-delhis-heat-action-plan-a-ground-study-part-i.
27. The Leaflet. Auditing Delhi's Heat Action Plan: A Ground Study, Part II (11 August 2026). https://theleaflet.in/governance-and-policy/auditing-delhis-heat-action-plan-a-ground-study-part-ii.
28. Ministry of Labour & Employment, Government of India. Ministry of Labour & Employment issues nationwide heatwave advisory to protect workers and labourers. Press Information Bureau (28 April 2026).
29. Sareen, R. & Singh, M. Making Delhi Heat-Resilient: A Roadmap with the Focus on Vulnerable Groups (Centre for Science and Environment, 2026).
30. California Code of Regulations. Title 8, section 3395, Heat Illness Prevention in Outdoor Places of Employment.
31. Ajuntament de Barcelona. More climate shelters to offer protection from the heat (4 June 2025). https://ajuntament.barcelona.cat/agenda2030/en/actualitat/creix-la-xarxa-de-refugis-climatics-per-protegir-se-de-la-calor-1524129.
32. Li, Z. et al. Machine learning early warning for urban heat risk with CMIP6 projections. J. Environ. Manage. 413, 130304 (2026). https://doi.org/10.1016/j.jenvman.2026.130304.
33. Kacker, K., Srivastava, P. & Mukherjee, M. Personalized heat stress early warning system for an urban area. Environ. Int. 199, 109507 (2025). https://doi.org/10.1016/j.envint.2025.109507.
34. World Health Organization. Ethics and Governance of Artificial Intelligence for Health: WHO Guidance (WHO, 2021).
35. Tabassi, E. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1 (National Institute of Standards and Technology, 2023). https://doi.org/10.6028/NIST.AI.100-1.
36. Williams, H. T. P. et al. Developing trustworthy AI for climate services at scale. npj Clim. Action 5, 74 (2026). https://doi.org/10.1038/s44168-026-00420-z.

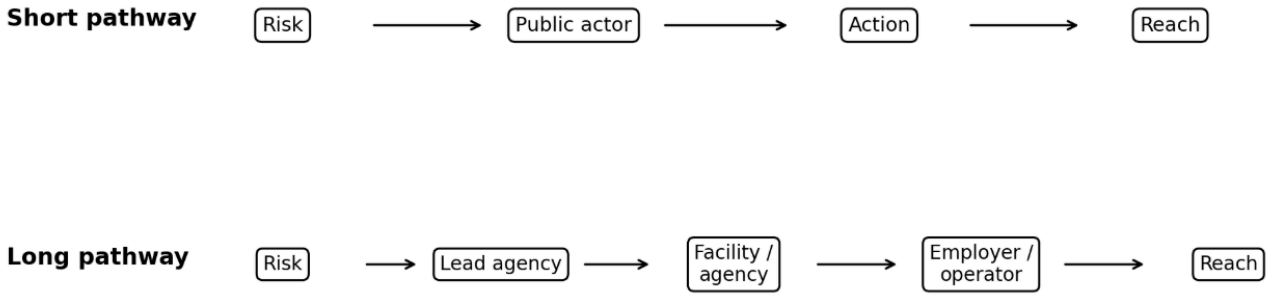
Figures



No stage is inferred from the stage before it. Missing downstream evidence remains unknown.

Figure 1

Risk-to-Protection Translation. The framework separates upstream risk intelligence from downstream protective states. No stage is inferred from the stage before it; missing downstream evidence remains unknown.



Each transfer can add delay or ambiguity; coordination, monitoring, incentives, and enforcement can compensate.

Figure 2

Handoff friction and compensating institutional controls. Longer pathways need not fail, but each additional transfer must be made observable and governable through coordination, monitoring, incentives or enforcement.

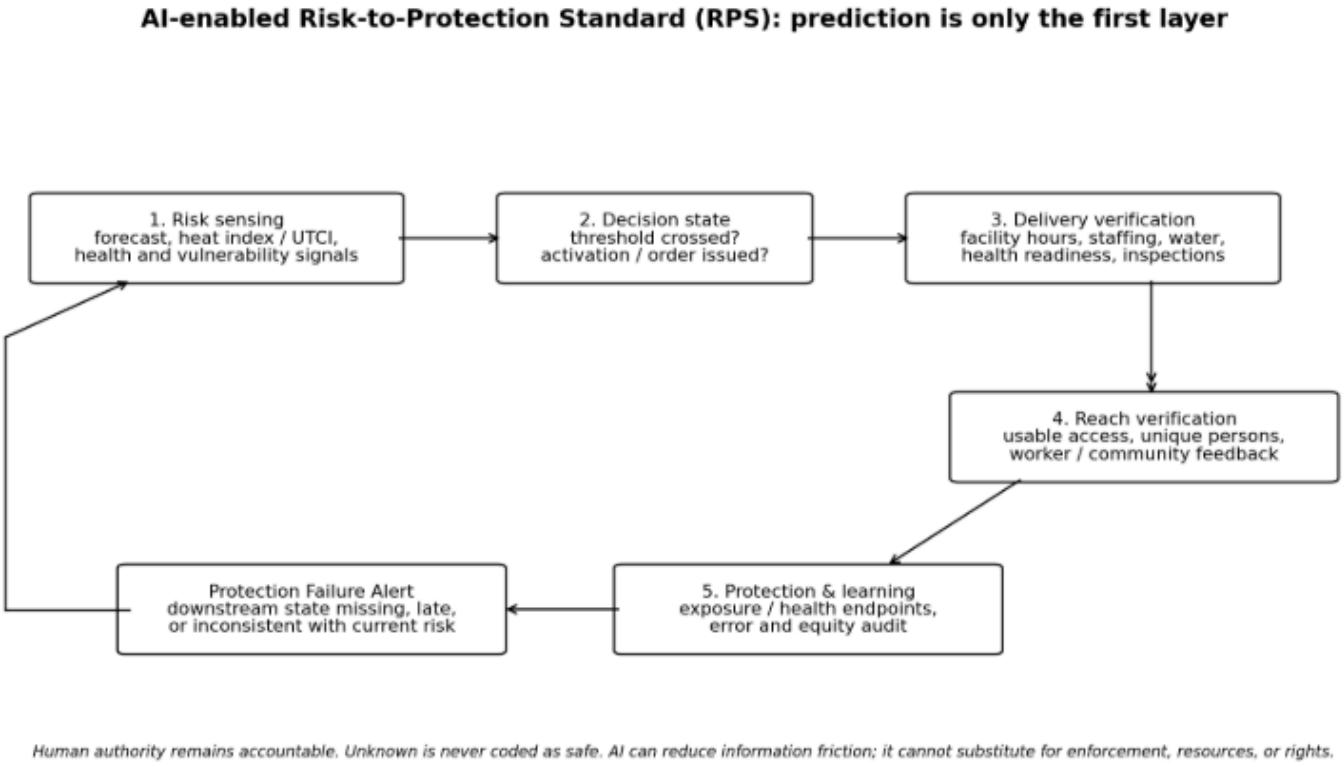


Figure 3

Proposed AI-enabled Risk-to-Protection Standard. AI connects risk intelligence to downstream verification and escalation while accountable human authority remains responsible for decisions and corrective action.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- NPJClimateActionSupplementaryInformation.pdf