A Report on "Mortality Caused by Tropical Cyclones in the United States" by Young and Hsiang (2024)

Reviewer 2

February 13, 2026

v1

isitcredible.com

Disclaimer

This report was generated by large language models, overseen by a human editor. It represents the honest opinion of The Catalogue of Errors Ltd, but its accuracy should be verified by a qualified expert. Comments can be made here. Any errors in the report will be corrected in future revisions.

> I am wiser than this person; for it is likely that neither of us knows anything fine and good, but he thinks he knows something when he does not know it, whereas I, just as I do not know, do not think I know, either. I seem, then, to be wiser than him in this small way, at least: that what I do not know, I do not think I know, either.
>
> Plato, *The Apology of Socrates*, 21d

> To err is human. All human knowledge is fallible and therefore uncertain. It follows that we must distinguish sharply between truth and certainty. That to err is human means not only that we must constantly struggle against error, but also that, even when we have taken the greatest care, we cannot be completely certain that we have not made a mistake.
>
> Karl Popper, 'Knowledge and the Shaping of Reality'

Overview

Citation: Young, R. and Hsiang, S. (2024). Mortality Caused by Tropical Cyclones in the United States. *Nature*. Vol 635, pp. 121–128.

URL: https://doi.org/10.1038/s41586-024-07945-5

Abstract Summary: This study evaluates the long-term effects of tropical cyclones on human mortality in the contiguous United States from 1930 to 2015, finding a robust and persistent increase in excess mortality that is substantially higher than official direct death counts.

Key Methodology: Econometric deconvolution of state-level all-cause mortality time series using a distributed lag model to estimate the impulse-response function of mortality to tropical cyclone wind speed incidence, accounting for various confounding factors.

Research Question: What is the full health impact, specifically the long-term effects on human mortality, of tropical cyclones in the contiguous United States?

Summary

Is It Credible?

Young and Hsiang present a provocative analysis suggesting that tropical cyclones (TCs) in the United States are a major, unrecognized driver of long-term mortality. Their headline claim is that between 1930 and 2015, TCs caused 3.6 to 5.7 million excess deaths, a figure that would rival the mortality burden of all infectious diseases or motor vehicle accidents over the same period (p. 127). This extraordinary estimate is derived not from counting immediate storm victims, but by using a "reduced-form" econometric model to detect a statistical signal of excess death that persists for 15 years (172 months) after a storm makes landfall (p. 121). While the authors employ rigorous placebo tests to validate their statistical model, the sheer magnitude of the claim and the extended duration of the effect strain plausibility in the absence of a clearly identified mechanism.

The credibility of these figures rests entirely on the validity of attributing deaths to a storm that occurred more than a decade prior. The study finds that the mortality response does not peak immediately, but rather rises slowly to reach a maximum approximately 68.6 months—nearly six years—after the event (p. 123). The authors acknowledge that they do not identify the underlying mechanisms driving this delayed peak, though they hypothesize economic disruption or social network decay as potential causes (p. 127). Without a concrete biological or social pathway to explain why mortality would crest six years post-disaster, the result risks being a statistical artifact of the "deconvolution" method used to separate overlapping storm signals. The reliance on accumulating small, statistically estimated monthly effects over such a long horizon makes the total burden highly sensitive to model specification and potential unobserved confounders.

Perhaps the most counter-intuitive finding concerns infant mortality. The authors

estimate that infants under one year of age face the highest risk, yet 99% of these excess infant deaths occur more than 21 months after the storm (p. 123). As the authors note, this implies that the infants were not even conceived when the storm hit. For this to be causal, the storm must degrade the local environment or health-care infrastructure so profoundly that it significantly increases infant mortality risk for children born years later. While not impossible, this asserts a level of long-term societal scarring that is difficult to reconcile with the generally rapid recovery of US infrastructure and the lack of specific cause-of-death evidence; indeed, the majority of excess deaths (58.9%) are attributed to a non-specific "other" category rather than identifiable causes like infectious disease (p. 124).

Methodologically, the reliance on state-level monthly average wind speeds as the primary measure of exposure introduces further uncertainty. By averaging wind speeds over an entire state for a full month, the specific, localized violence of a hurricane is diluted into a broad regional signal (p. 123). This aggregation makes it difficult to distinguish between genuine differential vulnerability and simple differential exposure. For instance, the finding that Black populations suffer significantly higher excess mortality (p. 124) is interpreted as evidence of greater vulnerability. However, if Black communities are systematically located in flood-prone areas within a state, they would experience higher physical exposure than the state-average wind metric captures. Furthermore, the model employs highly flexible eighth-order polynomials to control for trends (p. 127), a choice that risks overfitting the data and attributing residual low-frequency variation to the long-term effects of storms.

The Bottom Line

The article provides a statistically sophisticated argument that the official death toll of tropical cyclones is a vast underestimate, likely by orders of magnitude. However, the headline claim of millions of excess deaths relies on a causal link persisting

for 15 years without a clear mechanism, driven by a peak in mortality six years after the event. While the correlation is robust within the authors' model, the lack of a plausible biological or social explanation for such delayed effects suggests the true mortality burden may be overstated or conflated with other long-term regional stressors.

Potential Issues

Plausibility of the cumulative mortality burden: The article's headline conclusion is that tropical cyclones (TCs) caused between 3.6 and 5.7 million excess deaths in the contiguous U.S. from 1930 to 2015, a figure that would make TCs a previously unrecognized leading cause of death (pp. 125, 127). This extraordinary claim is derived by accumulating very small monthly mortality effects over a very long period of 14.3 years (172 months). The estimated monthly mortality increase at its peak is "0.042 deaths per 100,000" per 1 ms□1 of state-level wind speed incidence, occurring nearly six years after a storm (p. 123). While the authors conduct extensive placebo tests to ensure the statistical association is not an artifact (p. 123, Fig. 1g) and acknowledge that the magnitude of the finding is "surprising" (p. 127), the sheer scale of the result strains credulity. Deriving such a large total from the summation of small, delayed, and statistically estimated effects makes the final number highly sensitive to the model's assumptions and vulnerable to the possibility of unobserved confounding factors over long time horizons.

Lack of a plausible mechanism for the delayed mortality peak: The study's core empirical finding is a distinctive temporal pattern of excess mortality that rises slowly after a TC, peaks approximately 68.6 months (nearly six years) later, and then gradually declines over the subsequent eight years (p. 123). The article is explicitly a "reduced-form" analysis and states that it "does not identify the underlying mechanisms" that would explain this pattern (p. 127). While five broad hypotheses are proposed—including economic disruption, social network changes, and fiscal adjustments—the article does not offer a theory for why these or other factors would produce a mortality effect with this specific and highly delayed peak. Without a coherent biological or social mechanism to account for this temporal shape, the result is empirically adrift, making it difficult to distinguish from a potential statistical artifact.

Counter-intuitive findings on delayed infant mortality: The study reports that the highest mortality risk is for infants under one year of age. A particularly puzzling aspect of this finding is its timing: "Of these infant deaths, 99% occur more than 21 months after the TC," which the authors note indicates "that the infants were not conceived prior to landfall" (p. 123). This implies that the adverse post-disaster environment is so persistent and severe that it causes a substantial increase in mortality for babies born years, or even a decade, after the event. This requires a powerful and long-lasting causal pathway that specifically targets maternal or infant health well after the storm has passed. The authors note this finding is consistent with prior research on the long-term effects of disasters (p. 123), but the extreme nature of the delay and magnitude makes this result highly counter-intuitive and places a heavy burden on the statistical model's ability to isolate such a specific effect from other long-term trends in infant mortality.

Measurement of tropical cyclone exposure: The study's central causal variable—TC exposure—is measured as a monthly, state-wide average of maximum sustained wind speeds. This high level of spatial and temporal aggregation may create a significant disconnect between the proxy measure and the localized, intense nature of a TC event. By averaging wind speeds over the entire land area of a state for a full month, the signal of a powerful but geographically concentrated storm is substantially diluted. The authors acknowledge that wind speed is an "imperfect, proxy measure" and that their model does not explicitly account for storm surge or flooding (p. 123). They justify the choice based on data availability over their 86-year sample period and provide validation checks showing the metric is correlated with storm damages and rainfall. However, these correlations are modest, with the wind speed proxy explaining 36% of the variation in damages and 31% of the variation in rainfall at the state-by-storm level for a limited sample of storms (pp. 122, 127, Fig. 1d, 1e). This suggests the proxy may not adequately capture the destructive aspects of many storms, potentially introducing measurement error that could affect the accuracy of

the mortality estimates.

Tension in the causal identification strategy: The study's claim to causal inference relies on the premise of a "natural experiment" in which the timing of TCs is "as good as randomly assigned" to states over time (p. 127). This quasi-random assignment is essential for the econometric model to produce unbiased estimates. However, the article's own findings reveal that populations systematically adapt to their local TC risk, with states that experience TCs infrequently showing much higher vulnerability than states where TCs are common (p. 124). This finding of adaptation implies that the historical pattern of TC exposure has shaped the underlying characteristics of the populations being studied. While this does not invalidate the assumption that the timing of any given storm is random, it complicates the simple "natural experiment" framing of the study. The authors address this by employing standard econometric techniques, such as state fixed effects and modeling heterogeneous effects based on climate risk, to account for these pre-existing differences (pp. 123, 127).

Model specification and potential for overfitting: The study's econometric model is highly flexible, employing state-specific eighth-order polynomials to control for non-linear time trends in mortality, in addition to numerous other state- and time-specific fixed effects (p. 127). While the authors justify this complexity by demonstrating that simpler models fail placebo tests (p. 123), the use of such a high-order polynomial is not a parsimonious choice and risks overfitting the data. By removing so much variation through these flexible controls, the model searches for the TC signal in the small amount of residual variation that remains. This raises the possibility that the model could falsely attribute any low-frequency patterns in mortality not perfectly captured by the polynomials to the long-term effects of TCs.

Methodological assumptions of deconvolution: The study uses deconvolution to estimate a single, "characteristic mortality impulse-response function" that is assumed to be common to all TCs, scaled only by wind speed (p. 122). This approach relies on two key assumptions. First, it assumes additivity, meaning the combined

impact of multiple storms is the sum of their individual effects, which may not hold if a community is hit by a second storm while still recovering from a first. Second, it assumes a characteristic response shape. The authors state this is a "constraint" of the method (p. 127), but they do test for heterogeneity, finding no evidence that the response changed over time but allowing it to differ based on a state's historical climate risk (pp. 123, 128). Despite these tests, the core assumption represents a significant simplification of how diverse societies might respond to different types of disasters across nearly a century.

Conflation of vulnerability and exposure in disparities analysis: The analysis finds that Black individuals experience a cumulative excess mortality risk approximately 3.2 times higher than white individuals following a TC (p. 124). The study interprets this as evidence of differential vulnerability. However, the analysis relies on a state-level average wind speed as the measure of exposure for all subpopulations within that state. This design cannot distinguish between true differential vulnerability (a different health response to the same physical shock) and differential exposure (experiencing a more severe physical shock). The authors defend their interpretation by arguing there is "strong reason to believe that the spatial distribution of TC incidence within each state is orthogonal to the spatial distribution of underlying populations" (p. 127). However, if vulnerable populations are systematically located in areas with higher local risk within a state, such as flood-prone zones, they may experience a more severe storm impact than is captured by the state-level average, complicating the interpretation of the racial disparities.

Inconsistent findings on adaptation: The article presents two findings on adaptation that are potentially in tension. It finds strong evidence of adaptation across space, where states with a history of frequent TCs are less vulnerable than states where they are rare (p. 124). However, it finds "no evidence" of adaptation over time within the 1950–2015 sample period, meaning the mortality impact of a storm of a given intensity did not decrease over these 65 years (p. 123). It is surprising

that the process of adaptation would be evident in spatial patterns but apparently absent over a period of major advances in forecasting, medicine, and infrastructure. The authors acknowledge this contrast and hypothesize that the invisibility of the indirect excess deaths from TCs may have "prevented analogous adaptations" over time (p. 127).

Justification for the extended effect duration: The model estimates mortality effects that persist for 15 years or more, a choice the authors state was "partially motivated by analyses indicating that some economic effects of TCs persist for more than 15 years" (p. 122). Justifying a long-run mortality effect based on findings of long-run economic effects is a reasonable hypothesis but remains speculative. The choice of the lag length is a critical modeling decision, as a very long window increases the risk of capturing spurious low-frequency correlations. The authors validate their model and its long window using a series of placebo tests, which they report their model correctly passes (p. 123). However, the total mortality burden is sensitive to this choice of a 172-month integration window, and a more thorough sensitivity analysis of this parameter could strengthen the finding.

Interpretation of nonlinear effects: The study uses a range of models to estimate the total mortality burden, with the upper bound of 5.7 million deaths derived from a "Cubic Adaptation with 2 groups" model (p. 125, Extended Data Table 1). However, the Methods section notes that "these nonlinear effects are not themselves statistically significant" and that the linear and cubic models "only diverge (insignificantly) at very high levels of incidence" (p. 128). The authors justify using the cubic model because it contains information on how populations have adapted to their local climate (p. 128). Nonetheless, using a model with statistically insignificant nonlinear terms to define the upper bound of the article's headline finding is an aggressive interpretive choice that may overstate the certainty of the highest-end estimates.

Limited explanatory power of cause-of-death analysis: The analysis of cause-specific mortality finds that the largest share of TC-related excess deaths, 58.9%,

falls into a residual category of "'other' causes" (p. 124). This category is defined as the difference between all-cause mortality and the sum of five specified causes. The authors are transparent about this limitation, but attributing the majority of the effect to a non-specific, "catch-all" category provides little insight into the specific health pathways driving post-TC mortality.

Accounting for migration and spatial spillovers: The analysis uses states as the geographic unit of analysis, which may not fully account for population movements following major disasters. The authors acknowledge this as a limitation, stating their "estimates do not account for individuals who migrate outside of their state after a TC" (pp. 127–128). They argue that because their mortality rates are adjusted for population changes within the state, "TC-induced migration is too small to explain these findings," citing an unpublished PhD thesis (p. 128). However, if out-migration is selective (e.g., healthier individuals are more likely to leave), it could alter the mortality rate of the remaining population through a composition effect, confounding the estimated health impact of the storm.

Potential for confounding from downstream shocks: The study's reduced-form approach is designed to capture the total effect of a TC, including through indirect pathways like policy changes. The authors list "Fiscal adjustments" as a potential mechanism through which TCs affect health (p. 127). In this framework, a policy change caused by a TC is correctly considered part of its total effect. However, it is also possible that a major storm could be coincident with other long-term shocks (e.g., a regional economic downturn) that are not caused by the storm but influence both policy and mortality. The model's fixed effects and time trends are designed to control for such confounders, but disentangling these long-term, correlated processes is a significant analytical challenge.

Contextualization of excess deaths: To frame the significance of its findings, the article compares its estimate of 3.6–5.7 million TC-related deaths to deaths from "all motor vehicle accidents (2.0 million), infectious diseases (1.9 million) or US battle

deaths in wars (1.3 million)" during the study period (p. 127). This comparison equates statistically inferred "excess deaths"—an acceleration of mortality from various causes—with deaths that have a legally and medically certified direct cause. This is a rhetorical device that may be seen as an "apples-to-oranges" comparison, as the nature of the mortality being described is different.

Transparency in data harmonization: The analysis of cause-specific mortality required harmonizing data across the shift from ICD-9 to ICD-10 death certificate coding. The article states this was done by matching categories from a prior study by Barreca et al. (p. 127). However, it does not provide the specific mapping of codes used in the text, which limits the direct reproducibility of the cause-of-death analysis from the text alone. The authors state that the full data processing code is available, which implies the mapping is accessible in the provided replication files (p. 128).

Clerical and reporting issues: Several minor inconsistencies appear in the reporting of numerical results. First, the article reports different sample sizes and degrees of freedom in different sections without clear explanation for the discrepancies (pp. 122–123, 127). Second, there is a calculation error in the text: the reported mortality risk for infants (49.8 per 100,000) is 20 times higher than for adults aged 1–44 (2.49 per 100,000), not 16 times as stated (p. 123). Third, in the decomposition of long-term mortality trends, the reported percentage contribution from climatological factors (12%) appears to be an approximation or average of the two scenarios presented (+0.848 and +1.34 deaths per month) relative to the total trend (+9.2 deaths per month) (p. 127). These issues do not invalidate the study's conclusions but represent a lack of precision in the presentation of results.

Future Research

Identification of causal mechanisms: Future work should move beyond reduced-form estimates to identify the specific pathways driving delayed mortality. Research could link individual-level longitudinal health records with storm exposure data to determine if the "other" causes of death are driven by specific conditions such as stress-induced cardiovascular events, degradation of healthcare access, or economic deprivation, particularly explaining the six-year lag.

Refined exposure metrics: To address the limitations of state-level averaging, researchers should utilize high-resolution spatial data to measure exposure at the county or census-tract level. This would allow for a more precise distinction between wind damage, storm surge, and rainfall-induced flooding, and would help disentangle whether demographic disparities are driven by physiological vulnerability or simply by residence in higher-risk zones within a state.

Migration and displacement tracking: Given the potential for selective migration to confound state-level mortality rates, future studies should track individuals who move following a disaster. Using administrative data that follows populations across state lines would clarify whether the observed mortality effects are due to deteriorating conditions in the storm-hit area or the stress of displacement for those who leave.

Investigation of the infant mortality lag: The finding regarding infants born years after a storm requires targeted investigation. Research should focus on maternal health and birth outcomes in post-disaster years, specifically looking for evidence of maternal stress, nutritional deficits, or healthcare interruptions that could plausibly link a pre-conception disaster to post-birth mortality.

isitcredible.com