Post-Hurricane Vital Statistics Expose Fragility of Puerto Rico's Health System

Rolando J. Acosta[1] and Rafael A. Irizarry[1,2,*]

Abstract:

BACKGROUND: Hurricane Maria made landfall in Puerto Rico on September 20, 2017. As recently as May of this year (2018), the official death count was 64. This figure was criticized for being overly optimistic by groups with access to September and October demographic registry data. However, because the government was not making all post-hurricane mortality counts publicly available, fully understanding the hurricane's effect was challenging. A study describing a household survey, published on May 29, 2018, reported a much higher death count estimate, as well as evidence of population displacement, extensive loss of services, and a prolonged death rate increase lasting until the end of the survey period. Three days after this report was published, the government released death registry data. Here we use these data as well as data from other states to provide a detailed description of the effects on mortality of hurricane María. We compare these effects to those of other hurricanes.

METHODS: We fit a statistical model to mortality data that accounts for seasonal and non-hurricane related yearly effects. We then estimated the deviation from the expected death rate as a function of time using natural cubic splines that allowed discontinuities at hurricane landfall dates. We fit this model to 1985-2018 Puerto Rico daily data, which includes the dates of hurricanes Hugo, Georges, and María, 2015-2018 Florida daily data, which includes the dates of hurricane Irma, 2002-2004 Louisiana monthly data, which includes the date of hurricane Katrina, and 2000-2016 New Jersey monthly data, which includes the date of hurricane Sandy.

RESULTS: We estimated death rate increases on the day of the hurricane of 689%, 74%, 33%, 10%, and 2% for Katrina, María, Georges, Hugo, and Irma, respectively. No increase was noted for Sandy. We find a prolonged increase in death rate after María and Katrina, lasting at least 207 and 125 days, resulting in excess deaths estimates of 3,433 (95% CI, 3,189-3,676), and 1,832 (95% CI, 1,600-2064) respectively, showing that María had a more long term damaging impact. Surprisingly, we also find that in 1998, Georges had a comparable impact to Katrina's with a prolonged increase of 106 days resulting in 1,427 (95% CI, 1,247-1,607) excess deaths.

---

[1] Department of Biostatistics, Harvard T.H. Chan School of Public Health
[2] Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute
* rafa@jimmy.harvard.edu

For Hurricane María, we find sharp increases in a small number of causes of deaths, including diseases of the circulatory, endocrine and respiratory system, as well as bacterial infections and suicides.

CONCLUSIONS: Our analysis suggests that since at least 1998, Puerto Rico's health system has been in a precarious state. Without a substantial intervention, it appears that if hit with another strong hurricane, Puerto Ricans will suffer the unnecessary death of hundreds of its citizens.

Introduction

Hurricane Maria made landfall in Puerto Rico on September 20, 2017, interrupting the water supply, electricity, telecommunications networks, and access to medical care. In early May 2018, the official death count stood at 64[1]. This figure was in conflict with estimates obtained by groups that ostensibly were able to obtain death counts for September and October from the demographic registry of Puerto Rico. By comparing these two numbers to historical averages, additional deaths attributable to the hurricane were estimated to be in excess of 1,000[2-5]. However, as of May 2017, the government of Puerto Rico was not releasing the 2017-2018 data.

On May 29, 2018 a paper was published[6], here referred to as the *Harvard Study,* describing a survey of 3,299 households and reporting a death count estimate of 4,645 (95% CI, 793 to 8,498). It also reported an extensive loss of services after the hurricane: "On average, households went 84 days without electricity, 68 days without water, and 41 days without cellular telephone coverage after the hurricane and until December 31, 2017." Perhaps most importantly, the study showed evidence of a sustained effect on mortality throughout this extended period. These findings underscored the importance of a careful analysis to determine if, for example, there was a systematic increase in deaths due to indirect effects, if a specific demographic was at greater risk, and what type of medical conditions needed most attention.

The Harvard study received worldwide media coverage and three days after its publication, while under significant public pressure and facing a lawsuit[7], the government finally made the data public and acknowledged the possibility of a higher death count. Specifically, on June 13, the government updated the death count to 1,427 following the release of partial death records the day before[8]. This number is consistent with the value one obtains by simply subtracting the

2017 counts from the average for 2013 to 2016 using the first table made available by the government (Supplementary Table 1): (2928 - 2399) + (3040 - 2514) + (2671 - 2418) + (2820 - 2701) = 1427. Santos-Lozada and Howard[9] used the newly released data to update their previous estimate and reported an excess death estimate of 1,139 (95% CI, 1,006-1,272). However, this is a downwardly biased estimate because rather than subtracting the expected count, the authors subtract the upper limit of a 95% confidence interval for the expected count and they only include September, October and November data (we also find that the reported size of the confidence interval is too small and inconsistent with the data). Furthermore, neither of these estimates took into account the population displacement described by the Harvard study and others[10,11]. An analysis, posted online, took into account population displacement and showed data visualization suggesting a much larger count[12]. A government commissioned study released over two months later came to a similar conclusion and provided an estimate of 2,975 (95% CI: 2,658-3,290) for the total study period of September 2017 through February 2018[13].

Here we use these daily counts and individual level mortality data to provide a detailed and more accurate picture of the effect hurricane María had on mortality in Puerto Rico. We use data provided by Teralytics Inc. to estimate the population size. We compare the death rate increases to those observed in Hugo and Georges, two previous hurricanes in Puerto Rico, Katrina in Louisiana, Irma in Florida, and Sandy in New Jersey. We find a disturbing pattern in the Puerto Rico data.

Methods

Data

We set out to obtain detailed mortality and population size data related to hurricanes Hugo, Georges, and María in Puerto Rico, Katrina in Louisiana, Sandy in New Jersey, Harvey in Texas, and Irma in Florida. We requested individual death information but, as described below, this was not always available. We used whatever data was made available, which resulted in a mix of individual, daily and monthly data.

Hurricanes Hugo, Georges, and María (Puerto Rico)

We requested daily death count data from the Department of Health of Puerto Rico and obtained data from January 1985 to June 2015. We also requested individual level information with no personal identifiers from the Department of Health of Puerto Rico and obtained individual death records including date, gender, age, and up to ten causes of death from

January 2015 to June 2018. We used these data to construct the daily counts for the 2015-2018 period. Exploratory data analysis showed that data from after May 31, 2018 were incomplete (Supplementary Figure 1). We therefore discarded data past May 31, 2018. Yearly population estimates for the island were obtained from the Statistical Institute of Puerto Rico. We computed daily population estimates via linear interpolation (Supplementary Figure 2A). To obtain a more accurate estimate of the population displacement after Hurricane Maria, we used de-identified cellphone data to estimate population movement from the island to the United States and vice versa. The data was provided by tech company Teralytics Inc.[14] and spanned from September 2017 to May 2018 (Supplementary Figure 2B). We combined these two datasets to obtain a final estimate of the population of Puerto Rico for the period in question (Supplementary Figure 2C). Details on how we did this can be seen in our code.

Hurricane Irma (Florida)

We requested daily death counts from Florida's Vital Statistic System and obtained data from 2015 to 2018. For consistency, we discarded data past May 31, 2018. Yearly population estimates were obtained from the US Census for 2015-2017. We computed daily population estimates using interpolation. We extrapolated using a linear model to compute daily population estimates for 2018. Furthermore, we used data provided by Teralytics to estimate changes in population, which may be owing to people fleeing to other states due to Hurricane Irma, or Puerto Ricans migrating to the state due to Hurricane Maria (Supplementary Figure 2D).

Hurricane Katrina (Louisiana)

We requested daily death counts from Louisiana's Vital Statistic System and obtained data from 2003 to 2006. For Louisiana, we also obtained monthly death counts from the Underlying Cause of Death database through CDC WONDER for 2000-2008[15]. These two datasets did not match for the months following hurricane Katrina (Supplementary Figure 3). Since the data for August 2005 matched, we used the daily data to divide the monthly counts for August into before, during, and after the hurricane counts. We obtained population estimates from the US Census and computed daily population estimates via linear interpolation (Supplementary Figure 2E).

Hurricane Sandy (New Jersey)

We obtained monthly death counts for New Jersey from the Underlying Cause of Death database from 2000 to 2016. We also obtained yearly population estimates from the US Census and interpolated to obtain monthly estimates. (Supplementary Figure 2F)

Hurricane Harvey (Texas)

We were unable to obtain data related to Hurricane Harvey. We requested daily death counts from Texas' Vital Statistic System for a period including the dates Hurricane Harvey made landfall, but our petition was denied. The Underlying Cause of Death database does not have data available for 2017.

Statistical Methods: Daily counts

We assumed that the death counts $Y_{i,j}$ for the j-th day of the i-th year follow a Poisson distribution with rate:

$$\lambda_{i,j} = N_{i,j} exp\{\alpha_i + s(j) + f(t_{i,j})\}.$$

Here $N_{i,j}$ is an offset to account for the changing population size, α_i accounts for the year-to-year variability not due to hurricanes, $s(j)$ is the seasonal effect for the j-th day of the year, $t_{i,j} = 365 * (i - 1) + j$ is time in days, and $f(t_{i,j})$ accounts for the remaining variability not explained by the Poisson variability. So, for example, a virus epidemic will make $f(t_{i,j})$ go up slowly, eradication of this epidemic will make $f(t_{i,j})$ go down slowly, and a catastrophe will make $f(t_{i,j})$ jump up sharply. We therefore assume $f(t_{i,j})$ is a smooth function of $t_{i,j}$ except for the days hurricanes make landfall in which the function may be discontinuous.

Because $s(j)$ is seasonal, we use Fourier's theorem and model it as:

$$s(j) = \mu + \sum_{k=1}^{K} \{A_k cos(2\pi kj/365) + B_k sin(2\pi kj/365)\}$$

Note that we include an intercept μ, which represents the baseline rate for the entire period being studied. We assume that $f(t_{i,j})$ is a natural cubic spline with L equally spaced knots.$\tau_1, \dots, \tau_L$, except that the closest knot to the hurricane day is changed to be exactly on the hurricane date and we permit a discontinuity at this knot. Since natural cubic splines can be represented as a linear combination of basis function and $s(j)$ is a linear combination of known functions, ours is a generalized linear model (GLM) and, in theory, we can estimate α_i, s, and f using maximum likelihood estimates. However, because we want f to be flexible enough to

capture relatively high-frequency signals, such as those generated by virus epidemics, we instead implement a modular approach that first estimates the α_is and s, then uses these as offsets to estimate f. Specifically, we assume that for the non-hurricane years, $\alpha_i + f(t_{i,j})$ average out to 0 across years and use this assumption to estimate $s(j)$ using the standard GLM approach. We then use the estimated $\hat{s}(j)$ as an offset to estimate the year-to-year deviations α_i using only months not affected by hurricanes (March to August). We then use the estimate $\hat{\alpha}_i + \hat{s}(j)$ as an offset and estimate $f(t_{i,j})$ with the MLE and use standard GLM theory to estimate standard errors. Finally, due to lack of data to estimate α_i for 2018, we first obtained the estimates for the other years and then extrapolated (Supplementary Figure 4). See code for the details.

For the seasonal effect, we use *K=3* since it results in a smooth estimate that captures the general shape of the seasonal trend (Supplementary Figure 5). We used 4.5 knots per year to model $f(t_{i,j})$ as this results in a smooth estimate that captures the trend observed in the data (Supplementary Figure 6). Based on these plots, we removed years 2001 (Supplementary Figure 6D) and 2014 (Supplementary Figure 6F) for the computation of the seasonal effect: in 2001 there appears to be undercounts in January and in 2014 we see an increased mortality rate in agreement with the Chikungunya epidemic. Diagnostics plots for the residuals, after removing 2011 due to the undercounts, further show that the model fits the data (Supplementary Figure 7). Details of how we implemented this approach can be learned by studying our code included in this GitHub repository: (https://github.com/rafalab/maria).

Statistical Methods: Monthly counts

For the monthly data, we fit a monthly version of the model above. Because the counts are much larger once we aggregate at this level, we made use of the normal approximation to the Poisson. Specifically, we defined the monthly rates as $r_{i,m} \equiv \overline{Y}_{i,m}/\overline{N}_{i,m}$ where $\overline{Y}_{i,m}$ is the average number of deaths in month m of year i and $\overline{N}_{i,m}$ is the person years for that period. Thus, we collapsed the model above to a monthly version as follows: we assumed that the monthly rates can be described with the following model:

$$r_{i,m} = \alpha_i + \overline{s}_m + \beta_{i,m} X_{i,m} + \varepsilon_{i,m}.$$

Here α_i accounts for year-to-year variability as in daily data model and $\overline{s}_m$ is the average of $s(j)$ for all days j in month m. Because we no longer need splines to model the effects, we instead use indicator functions to denote if a month/year was affected by the hurricane. Specifically, we define $X_{i,m}$ as an indicator that is 1 for the months m in year i that were affected by the hurricane and 0 otherwise. The parameter $\beta_{i,m}$ thus represents the effect of the hurricane on death rate and is equivalent to the integral of $f(t_{i,j})$ for $t_{i,j}$ in month m of year i. The natural, yet non-hurricane related variability, is represented by the term $\varepsilon_{i,m}$ which are assumed to be independent and normally distributed with average 0 and month-specific standard deviation σ_m. Notice that this is a standard linear model and the estimates can be obtained with least squares.

Excess death estimates

The first step in estimating excess deaths was to determine the period of indirect effect of the hurricane. We define this period as the interval starting on the day the hurricane made landfall, denoted here with t_0, until the first day, $t_{i,j} > t_0$, for which there is no longer a positive increase: $f(t_{i,j}) \leq 0$. Because we do not observe $f(t_{i,j})$, and instead obtain an estimate denoted here with $\hat{f}(t_{i,j})$, we take the conservative approach of defining the indirect effect period with the day $t_{i,j} > t_0$ for which the lower part of a marginal 95% confidence interval is 0: $\hat{f}(t_{i,j}) - 1.96\widehat{SE}\{\hat{f}(t_{i,j})\} \leq 0$. Once we have this interval, denoted here with I, we define the excess deaths by adding the observed deaths minus expected deaths for every day in the interval:

$$\text{Excess deaths} = \sum_{t_{i,j} \in I} \left[Y_{i,j} - N_{i,j} \exp\{\hat{\alpha}_0 + \hat{s}(j)\} \right]$$

We construct a 95% confidence interval using the following approximation for the Poisson model:

$$\text{Var}(\text{Excess deaths}) = \sum_{t_{i,j} \in I} N_{i,j} \exp\{\hat{\alpha}_0 + \hat{s}(j)\}$$

Natural variability

The standard error computed above does not take into account the natural variability accounted for by $f(t_{i,j})$ in non-hurricane years. As mentioned above, this quantity represents natural variability not accounted for by the Poisson variability. We therefore quantify the day-specific

variability with the observed standard deviation across years for $f(t_{i,j})$. We refer to this as the *natural variability.*

Cause of Death

To examine if any cause of death was more prevalent after the hurricane, we used the individual records data spanning 2015-2017. We did not include 2018 data because it appears that the cause of death data is incomplete for the time after December 31, 2017 (Supplementary Figure 8). We divided causes of death into 30 categories (Supplementary Table 2) and, for each of these, we computed the observed death rate during the September 20 - December 31 period for 2017, and compared to the expected rates computed with the 2015-2016 data for that same period. We used the Poisson model to compute confidence intervals for these quantities. To estimate a daily effect for an ICD group, we fit the Poisson GLM model described above to the daily counts for a given ICD group.

Results

Indirect effects

The death rate in Puerto Rico increased by 73.9% (95% CI, 63.1%-85.4%) the day after hurricane María made landfall (Figure 1, Supplementary Figure 6G). But the death rate did not return to historical levels until at least April 15, 2018. During the September 20, 2017 to April 15, 2018 period, the average increase in death rate was 22% (Figure 1, Supplementary Figure 6G). The effects of Katrina were much more direct. On August 29, 2005, the day the levees broke, there were 834 deaths, which translates to an increased in death rate of 689%. However, the increase in mortality rates for the four months following this catastrophe were 17%, 9%, 11%, and 5% percent, respectively (Supplementary Figure 9), substantially lower than for María. For Georges, a hurricane not considered to have had catastrophic effects, we observed a similar pattern to María: a sharp increase to 33% (95% CI, 25%-41%) on landfall day and a death rate not returning to historical levels until January 5, 1999 (Figure 1, Supplementary Figure 6C). The average percent increase in this period was 9%. None of the other hurricanes examined had noticeable indirect effects.

Excess deaths

We estimated excess deaths of 3,433 (95% CI, 3,189-3,676) for María in Puerto Rico, 1,832 (95% CI, 1,600-2,064) for Katrina in Louisiana, and 1,427 (95% CI, 1,247-1,607) for Georges in Puerto Rico (Figure 2). These estimates were calculated over periods of 207, 125, and 106

days after landfall for María, Katrina, and Georges, respectively. We note that the way in which these deaths accumulated through time were distinctively different. Namely, 39.9% of the excess deaths associated with Katrina occurred on Aug 29, 2005, the day the Levee's broke, while for the Puerto Rico hurricanes, the excess deaths accumulated slowly through a period of months after landfall (Figure 1, Supplementary Figure 6C, Supplementary Figure 6G).

Cause of Death

Increases in rates were not uniformly seen across all cases of death after hurricane María. Instead, we observed increases for a subset of the causes (Figure 3A). Not surprisingly, storm related deaths showed the largest increase. Although deaths directly related to the natural disaster increased the most, in terms of total excess deaths diseases of the circulatory, nervous, endocrine and respiratory systems explained well over 65% of deaths until at least December 31, 2017 (Table 1, Figure 3B). We note that these categories include heart and diabetes related problems. When examining the increase in death rate as function of time for these causes of death, we note that the indirect effects were substantial (Figure 4).

Death Rate by Age and Gender

The most affected demographic groups were individuals 80 years and older, followed by individuals between 70 to 79, and then individuals between 60 and 69 (Figure 5). Individuals between 50 and 59 seem to be mostly affected by indirect effects (Figure 5). Although younger demographics (< 49 years old) were not significantly affected (Supplementary Figure 10), these results demonstrate that a large proportion of the population was indeed affected by the indirect effects. Furthermore, there were no significant differences between genders (Supplementary Figure 10).

Natural variability in excess death rates

There has been much interest in estimating the excess deaths caused by hurricane María. Several confidence intervals have been reported in the literature based on the registry data. For example [605-1,039] for September and October [4], [1,006-1,273] for September to December [9], [2,658-3,290] for September to February [13], and we estimate [3,189-3,676] for September 20 to April 15. It is important to keep in mind that these are confidence intervals for the expected count assuming that in the periods after the hurricane, deviations of $f(t_{i,j})$ from 0 are entirely due to the hurricane. However, discussions around these numbers need to take into account the natural variability in $f(t_{i,j})$. Specifically, note that if we focus on years with no hurricanes, when

the expected excess death counts are 0, we still see deviations from the seasonal and yearly effects not accounted for by Poisson variability (Supplementary Figure 11A). In fact, the standard deviation of $\hat{f}(t_{i,j})$ taken across years was as high as 6% in parts of the winter (Supplementary Figure 11A). As a result, for a period of, for example, 103 days (September 20 to December 31) these levels of variability translate into a standard deviation of excess deaths of about 600 (Supplementary Figure 11B). We also note that this underscores the importance of going beyond analyses based on just monthly counts and historical averages. In our analyses we provided evidence of a hurricane effect by estimating and examining the shape of the effect as a function of time (Figure 1) and by studying specific causes of the excess deaths (Figure 3, Table 1).

Discussion

On August 29, 2005, a surge due to hurricane Katrina breached the levees of the Mississippi River-Gulf Outlet Canal and flooded several residential areas in the New Orleans area. This turned out to be catastrophic and appears to have caused over 600 direct deaths and up to 1,500 indirect deaths in the following four months. In a June 2006 report on the disaster[16] the U.S. Army Corps of Engineers admitted that faulty design specifications, incomplete sections, and substandard construction of levee segments, contributed to the damage and $14.5 billion[17,18] has been invested in befuddling stronger levees[18]. On September 21, 1998, Hurricane Georges made landfall in Puerto Rico causing great damage to an already fragile electrical grid. Simply plotting the raw mortality data for 1998 shows a disturbing increase in mortality rates (Supplementary Figure 12). Using a formal statistical model, we estimated 1,427 excess deaths due to this hurricane, an overall impact similar to that of Katrina in Louisiana. In contrast to the response to Katrina, as far as we know, no systematic effort was put in place to improve Puerto Rico's electrical grid nor its fragile health system. On the contrary, negligence and abandonment seems to have permitted the electrical grid to continue to deteriorate for the next 19 years[19]. Tragically, after hurricane María made landfall in Puerto on September 20, 2017, the electrical grid was destroyed leaving 100% of the population[20], including health facilities, without electricity. It has been well documented that restoration of the electrical grid has been slow[21], with some estimates reporting that only 30% of the population had electricity a month after the tropical storm[22]. We estimate that, as a result of this fragile health system, a large proportion of the population was affected and as many as 3,000 excess deaths occurred. The insights presented in our analysis should be considered in preparation efforts for the next hurricane.

Note that for in our analysis the number of knots used to estimate the daily variability effects, the number of harmonics used to estimate the seasonal effect and the way we extrapolate to estimate the year offset for 2018 were chose by visual inspection. Although not substantially, changing these parameters do change the final results of our analysis. As mentioned above, our analysis, including the code to recreate all the figures and tables is available at: https://github.com/rafalab/maria. We invite others to use our publically available code and data to try out other approaches.

Acknowledgments

We thank María M. Juiz Gallego and Jose A. López Rodríguez from the Department of Health of Puerto Rico for diligently providing all the data we requested. We thank the Puerto Rico Institute of Statistics for providing population data. We thank Canay Deniz, Andrea Samdahl, Lara Montini and Ilya Vasilenko from Teralytics for sharing their data and providing helpful explanations. We thank Matthew Kiang for many valuable demography insights and Deepak Lamba Nieves for suggesting readings on Puerto Rico's electrical grid. Finally, we thank all the authors of NEJM paper whose publication appears to have resulted in the government releasing the data. In particular, we thank Caroline Buckee for getting it all started.

References

1. Hoyos J. 8 months after hurricane maria, the death toll in puerto rico remains a mystery. May 20, 2018; https://abcnews.go.com/US/months-hurricane-maria-death-toll-puerto-rico-remains/story?id=55308995.
2. Santos A, Howard JT. Estimates of excess deaths in Puerto Rico following Hurricane Maria. 2017.
3. Robles F DK, Fink S, Almukhtar S,. Official Toll in Puerto Rico: 64. Actual Deaths May Be 1,052. December 9, 2017; https://www.nytimes.com/interactive/2017/12/08/us/puerto-rico-hurricane-maria-death-toll.html.
4. Rivera R, Rolke W. Estimating the death toll of Hurricane Maria. *Significance.* 2018;15(1):8-9.
5. Pascual OS. Nearly 1,000 more people died in puerto rico after hurricane maría. December 7, 2017; http://periodismoinvestigativo.com/2017/12/nearly-1000-more-people-died-in-puerto-rico-after-hurricane-maria/#.
6. Kishore N, Marqués D, Mahmud A, et al. Mortality in Puerto Rico after Hurricane Maria. *New England journal of medicine.* 2018.
7. Acevedo N. Puerto rico faces lawsuits over hurricane death count data. June 1, 2018; https://www.nbcnews.com/storyline/puerto-rico-crisis/puerto-rico-faces-lawsuits-over-hurricane-death-count-data-n879341.
8. Cheslow D. Puerto rico releases data on hundreds of deaths following hurricane maria. June 13, 2018; https://www.npr.org/2018/06/13/619440877/puerto-rico-releases-data-on-hundreds-of-deaths-following-hurricane-maria.
9. Santos-Lozada AR, Howard JT. Use of Death Counts From Vital Statistics to Calculate Excess Deaths in Puerto Rico Following Hurricane Maria. *Jama.* 2018.
10. Echenique M ML. Mapping puerto rico's hurricane migration with mobile phone data. May 11, 2018; https://www.citylab.com/environment/2018/05/watch-puerto-ricos-hurricane-migration-via-mobile-phone-data/559889/.
11. Meléndez E. Estimates of post-hurricane María exodus from Puerto Rico.
12. Irizarry RA. Estimating mortality rates in puerto rico after hurricane maría using newly released official death counts. 2018; https://simplystatistics.org/2018/06/08/a-first-look-at-recently-released-official-puerto-rico-death-count-data/.
13. Andrade E BN, Colon-Ramos U, Edberg M, Garcia-Meza A, Goldman A, Goldman L, Roess A, Sandberg J, Santos-Burgoa C, Zeger S, . *Ascertainment of the estimated excess mortality from hurricane maría in puerto rico.* The George Washington University;August 28, 2018.
14. Teralytics. https://www.teralytics.net.
15. Centers for Disease Control and Prevention NCfHS. Underlying cause of death. https://wonder.cdc.gov/ucd-icd10.html.
16. Link LE JJ, Stevenson J, Stroupe W, Mosher RL, Martin D, Garster JK, Zilkoski DB, Ebersole BA, Westerink JJ, Resio DT, Dean RG, Sharp MK, Steedman RS, Duncan JM, Moentenich BL, Howard B, Harris J, Fitzgerald S, Moser D, Canning P, Foster J, Muller B, . *Performance evaluation of the new orleans and southeast louisiana hurricane protection system.* June 1, 2006.
17. Kates RW, Colten CE, Laska S, Leatherman SP. Reconstruction of New Orleans after Hurricane Katrina: a research perspective. *Proc Natl Acad Sci U S A.* 2006;103(40):14653-14660.

18. Adelson J. New orleans levees rated 'high risk'; here's what drives designation, despite system's high marks. July 26, 2018; https://www.theadvocate.com/new_orleans/news/article_d5b79d10-90d6-11e8-8aaf-87b7bc7799f6.html.
19. DiSavino S. Puerto rico power grid faces generational threat in hurricane maria. September 19, 2017; https://www.reuters.com/article/us-storm-maria-puertorico-power/puerto-rico-power-grid-faces-generational-threat-in-hurricane-maria-idUSKCN1BU2HV?il=0.
20. Ruiz-Kulian G. El 100% de puerto rico está sin luz. 2017; https://www.elnuevodia.com/noticias/locales/nota/el100depuertoricoestasinluz-2359210/.
21. Shermeyer J. Assessment of electrical and infrastructure recovery in Puerto Rico following hurricane Maria using a multisource time series of satellite imagery. *arXiv preprint arXiv:180705854.* 2018.
22. *Hurricanes nate, maria, irma, and harvey situation reports.* October 30, 2017.

Figures & Tables

Table 1: Excess deaths by cause of death for categories with p-values below 0.05/30

Description	Observed	Expected	Excess	CI
Diseases of the circulatory system	2511.00	1808	703	[2389-2639]
Endocrine diseases	1118.00	788	330	[1037-1205]
Diseases of the nervous system	961.00	672	289	[886-1042]
Diseases of the respiratory system	898.00	629	269	[826-976]
Bacterial infections, intestinal infectious, and STDs	362.00	201	161	[315-416]
External causes of morbidity and mortality	632.00	498	134	[573-697]
Ill-defined and unknown causes of mortality	268.00	147	121	[228-315]
Diseases of the genitourinary system: urinary system	435.00	331	104	[386-490]
Metabolic diseases	174.00	106	68	[143-211]
Diseases of the skin and subcutaneous tissue	102.00	53	49	[78-133]
Mental and behavioural disorders	167.00	119	48	[138-203]
Storm	34.00	0	34	[5-248]
Suicide	81.00	48	33	[61-108]
Diseases of the musculoskeletal system and connective tissue	78.00	46	32	[58-105]
Leptospirosis	21.00	4	17	[10-45]

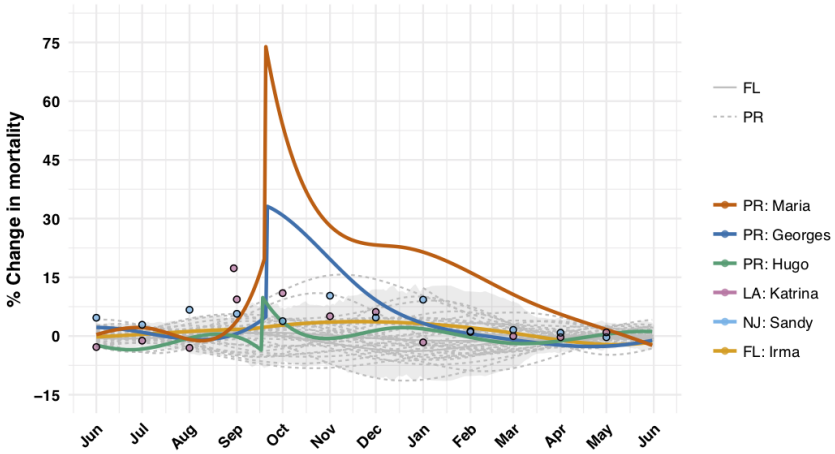

Figure 1: Estimated percent increases in death rates after hurricanes Hugo, Georges, and María in Puerto Rico, Katrina in Louisiana, Sandy in New Jersey and Irma in Florida. Estimates for Puerto Rico and Florida were obtained with daily data, while estimates for Louisiana and New Jersey were obtained with monthly data. The percent increase for Katrina on August 29, 2005, which is 689%, is outside the limits of the y-axis. The grey curves are percent increase estimates for non-hurricane years in Puerto Rico and Florida. The grey ribbon represents the range of variation (plus or minus two standard deviations) in death rate increases seen across 30 non-hurricane years in Puerto Rico. The standard deviations were calculated as daily median absolute deviance for the grey dashed curves

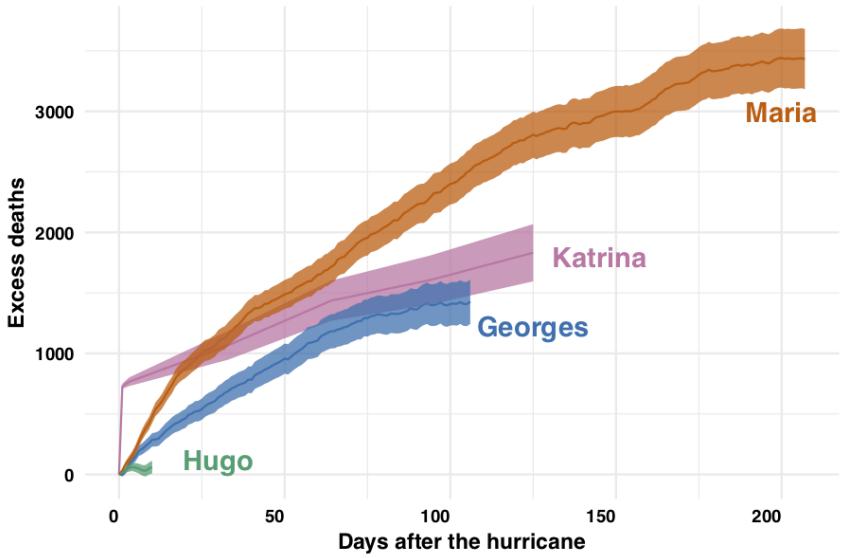

Figure 2: Excess deaths versus days after the hurricane for hurricanes Katrina, Georges, María and Hugo. The shaded areas represent marginal 95% confidence intervals for these estimates. The excess deaths are calculated for the periods affected by indirect effects

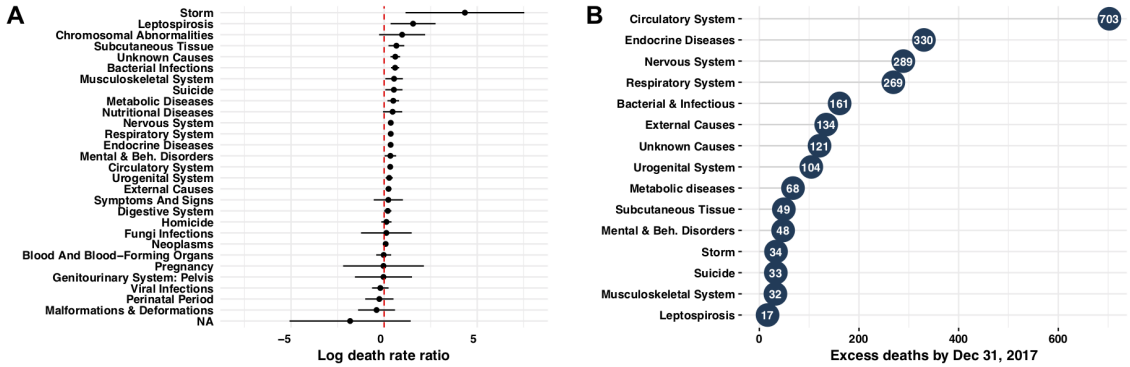

Figure 3: Increase in death rates for 30 causes of death after Hurricane María for the period starting September 20, 2017 and ending on December 31, 2017. A) Log death rate ratio estimates and Bonferroni corrected 95% confidence intervals. B) Estimated excess death by cause of deaths for causes of death with p-values below 0.05 / 30

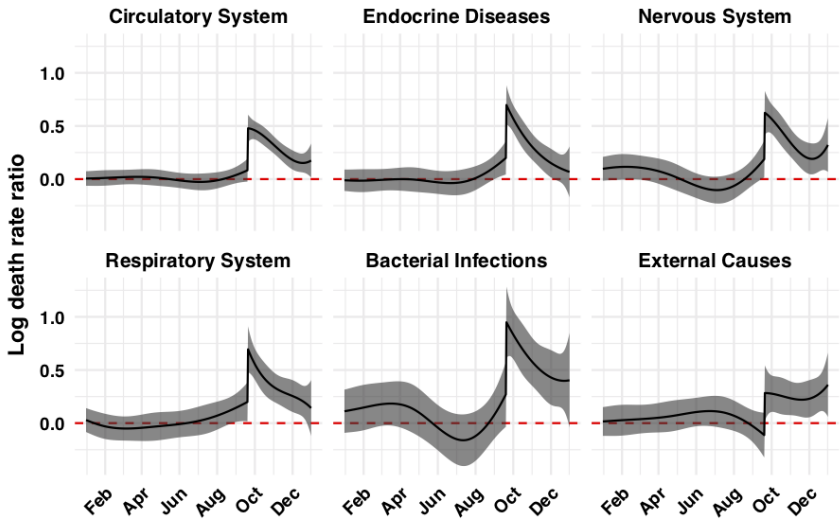

Figure 4: Estimated daily percent increase in death rate after hurricane María for the top six causes of excess deaths: Circulatory System, Endocrine Diseases, Nervous System, Respiratory System, Bacterial Infections, and External Causes. The gray ribbon represents marginal 95% confidence intervals

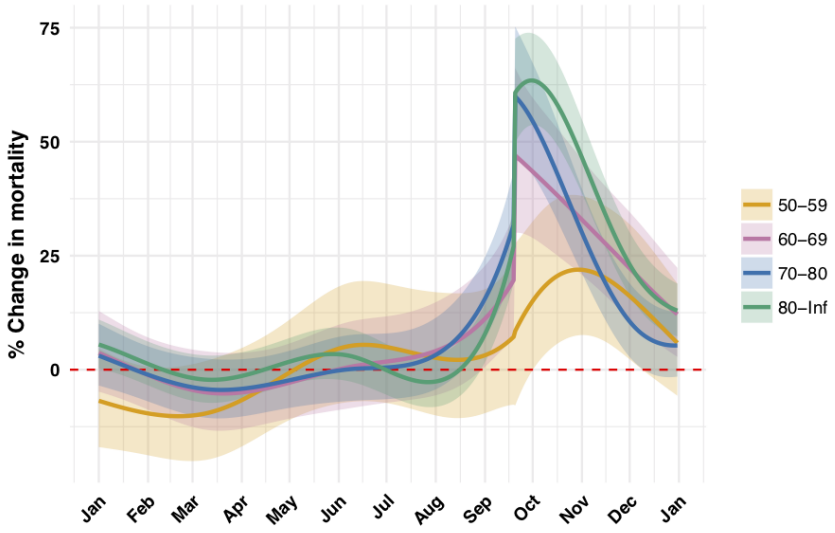

Figure 5: Estimated daily percent increase in death rate by age groups 60-69, 70-79, 80 and higher

Supplementary Material

Table 1: Data used by the government of Puerto Rico for the June 13 estimate of 1,427 excess deaths

Mes	2013	2014	2015	2016	2017	2018
1.00	2585.00	2484.00	2744.00	2743.00	2894.00	2821.00
2.00	2381.00	2216.00	2403.00	2592.00	2315.00	2448.00
3.00	2714.00	2489.00	2427.00	2458.00	2494.00	2647.00
4.00	2319.00	2396.00	2259.00	2241.00	2392.00	2223.00
5.00	2403.00	2479.00	2340.00	2312.00	2390.00	2226.00
6.00	2342.00	2389.00	2145.00	2356.00	2369.00	
7.00	2491.00	2423.00	2383.00	2457.00	2367.00	
8.00	2509.00	2539.00	2272.00	2427.00	2321.00	
9.00	2478.00	2495.00	2258.00	2368.00	2928.00	
10.00	2411.00	2894.00	2394.00	2357.00	3040.00	
11.00	2272.00	2651.00	2268.00	2484.00	2671.00	
12.00	2541.00	2891.00	2518.00	2854.00	2820.00	

Table 2: ICD 10 code map

ICD	Description
[A00,A79]	Bacterial infections, and other intestinal infectious diseases, and STDs
[A80,B34]	Viral infections
[B35,B89]	Infections caused by fungi, protozoans, worms, and infestations
[B90,B99]	Sequelae, and diseases classified elsewhere
[C00,D48]	Neoplasms
[D50,D89]	Blood and blood-forming organs and certain disorders involving the immune mechanism
[E00,E35]	Endocrine diseases
[E40,E68]	Nutritional diseases
[E70,E90]	Metabolic diseases
[F00,F99]	Mental and behavioural disorders
[G00,G99]	Diseases of the nervous system
[H00,H59]	Diseases of the eye and adnexa
[H60,H99]	Diseases of the ear and mastoid process
[I00,I99]	Diseases of the circulatory system
[J00,J99]	Diseases of the respiratory system
[K00,K93]	Diseases of the digestive system
[L00,L99]	Diseases of the skin and subcutaneous tissue
[M00,M99]	Diseases of the musculoskeletal system and connective tissue
[N00,N39]	Diseases of the genitourinary system: urinary system
[N40,N99]	Diseases of the genitourinary system: pelvis, genitals and breasts
[O00,O99]	Pregnancy, childbirth and the puerperium
[P00,P96]	Certain conditions originating in the perinatal period
[Q00,Q89]	Congenital malformations and deformations
[Q90,Q99]	Chromosomal abnormalities, not elsewhere classified
[R00,R69]	Symptoms and signs
[R70,R94]	Abnormal clinical and laboratory findings, not elsewhere classified
[R95,R99]	Ill-defined and unknown causes of mortality
[S00,T14]	Injury
[T15,T98]	Poisoning and certain other consequences of external causes
[V01,Y98]	External causes of morbidity and mortality

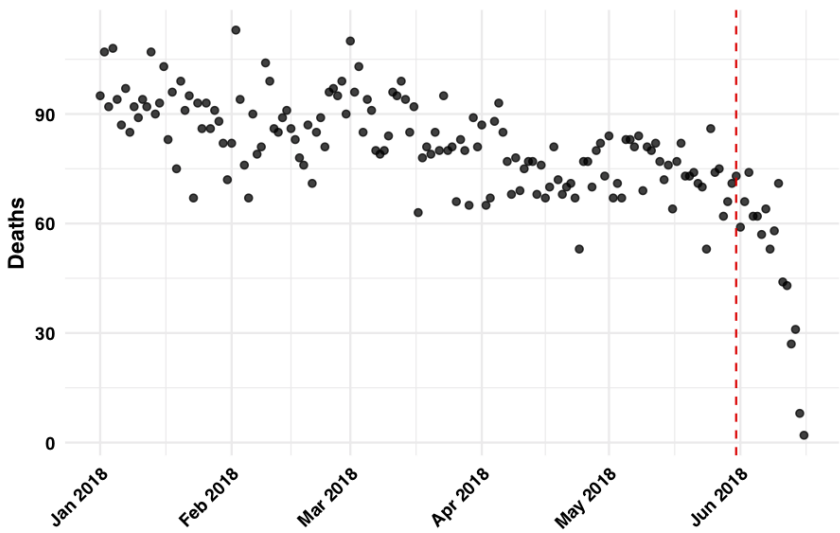

Figure 1: Reported deaths by day for Puerto Rico daily data in 2018. We excluded data to the right of the vertical dashed lines

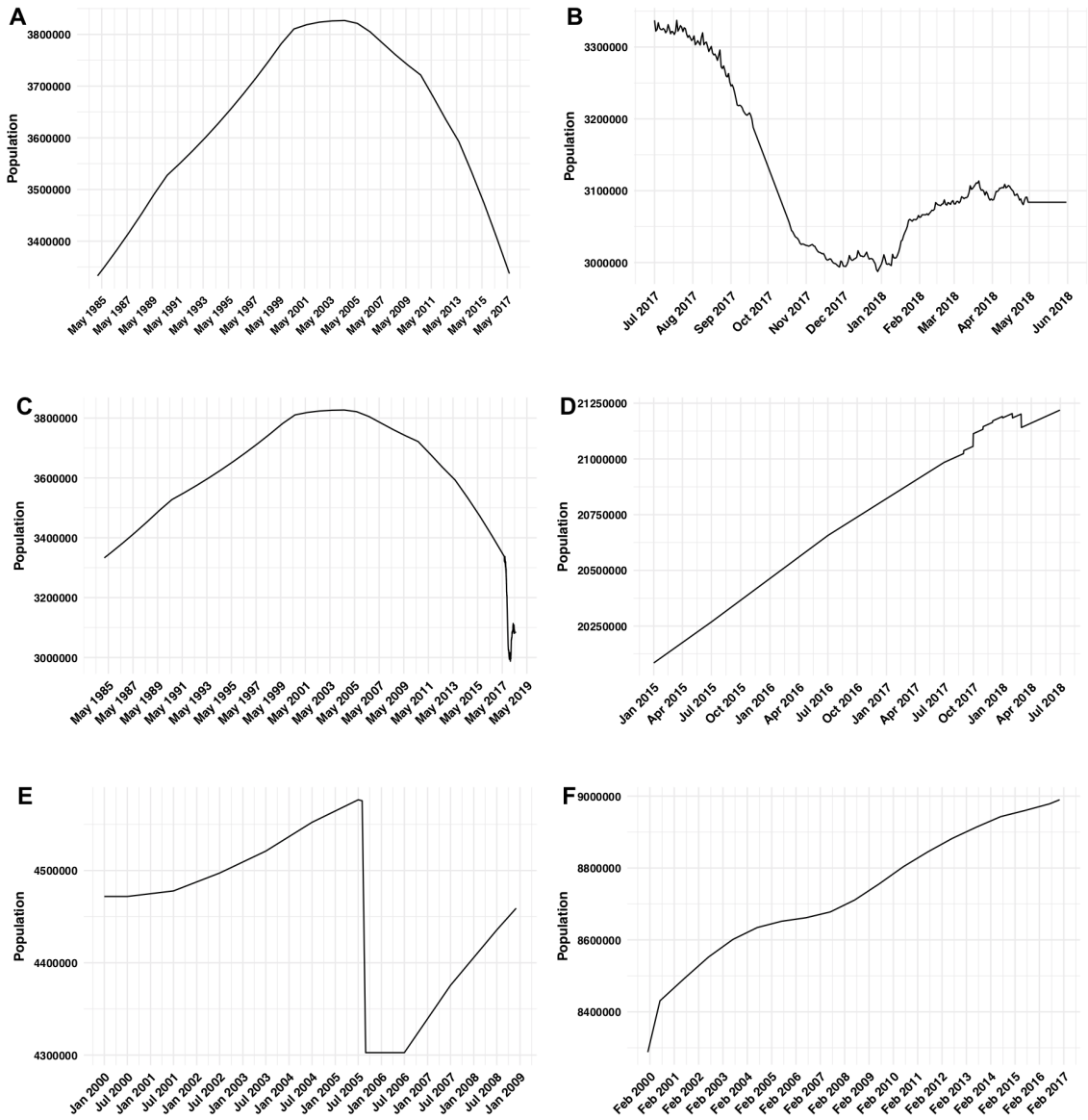

Figure 2: Estimated population sizes for six states. A) Interpolation of yearly estimates for Puerto Rico from 1985 to 2017. B) Estimates for Puerto Rico population from July 1, 2017 to June 1, 2018 based on proportion of population decreased provided by Teralytics Inc. C) Final estimate for Puerto Rico population. D) Estimated population for Florida. Includes increases based on Teralytics inc. estimate of immigration from Puerto Rico. E) Estimated population for Louisiana based on interpolation and extrapoliation. F) Estimate population for New Jersey based on interpolation

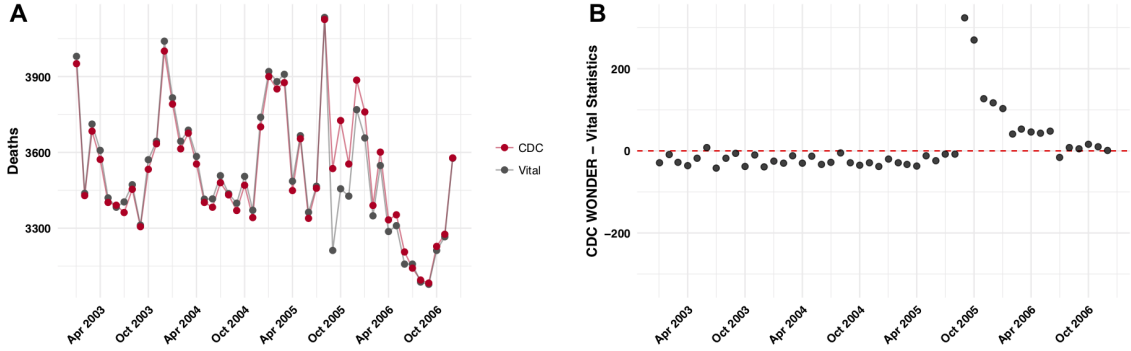

Figure 3: Comparison of census monthly death counts to Vital Statistics System monthly death counts. A) Total death counts. B) Difference between estimates

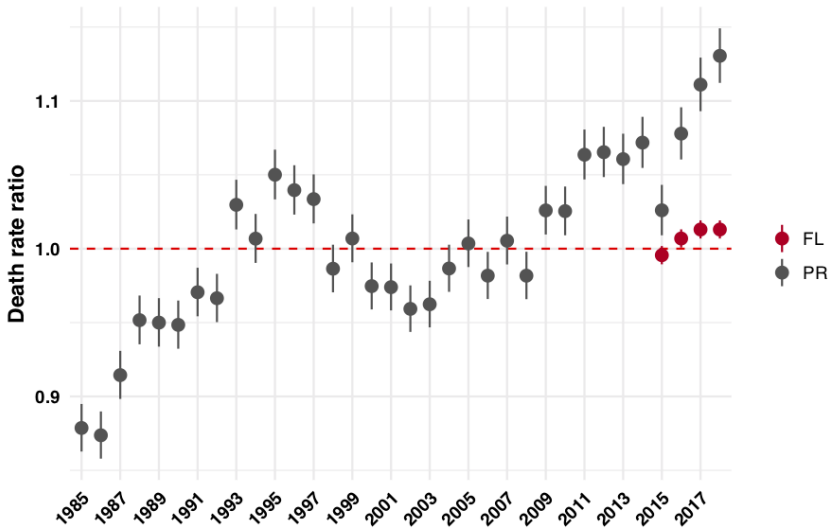

Figure 4: Estimated yearly offsets and marginal 95% confidence intervals for Puerto Rico and Florida obtained from fitting the GLM. The 2018 data is based on extrapolation obtained by fitting a natural cubic spline to the rest of the data for Puerto Rico and by assuming the same value as 2017 for Florida

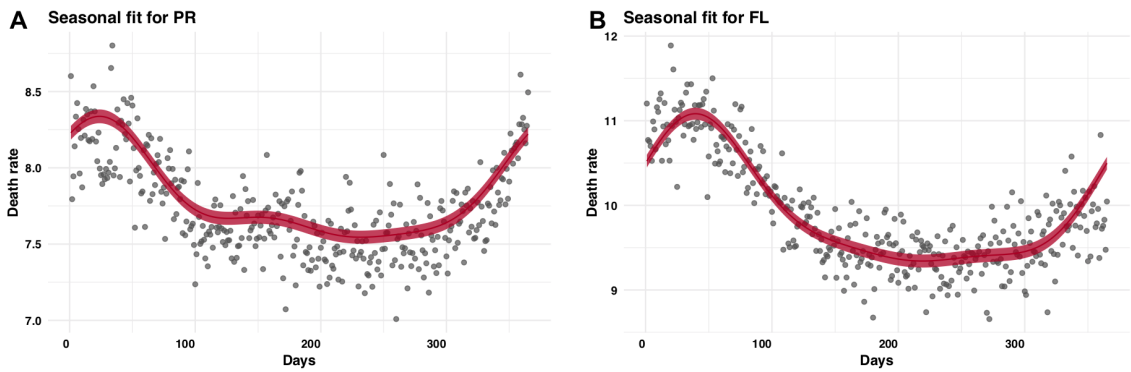

Figure 5: Estimated seasonal effects from the GLM and marginal 95% confidence intervals. A) Estimates for Puerto Rico. B) Estimates for Florida

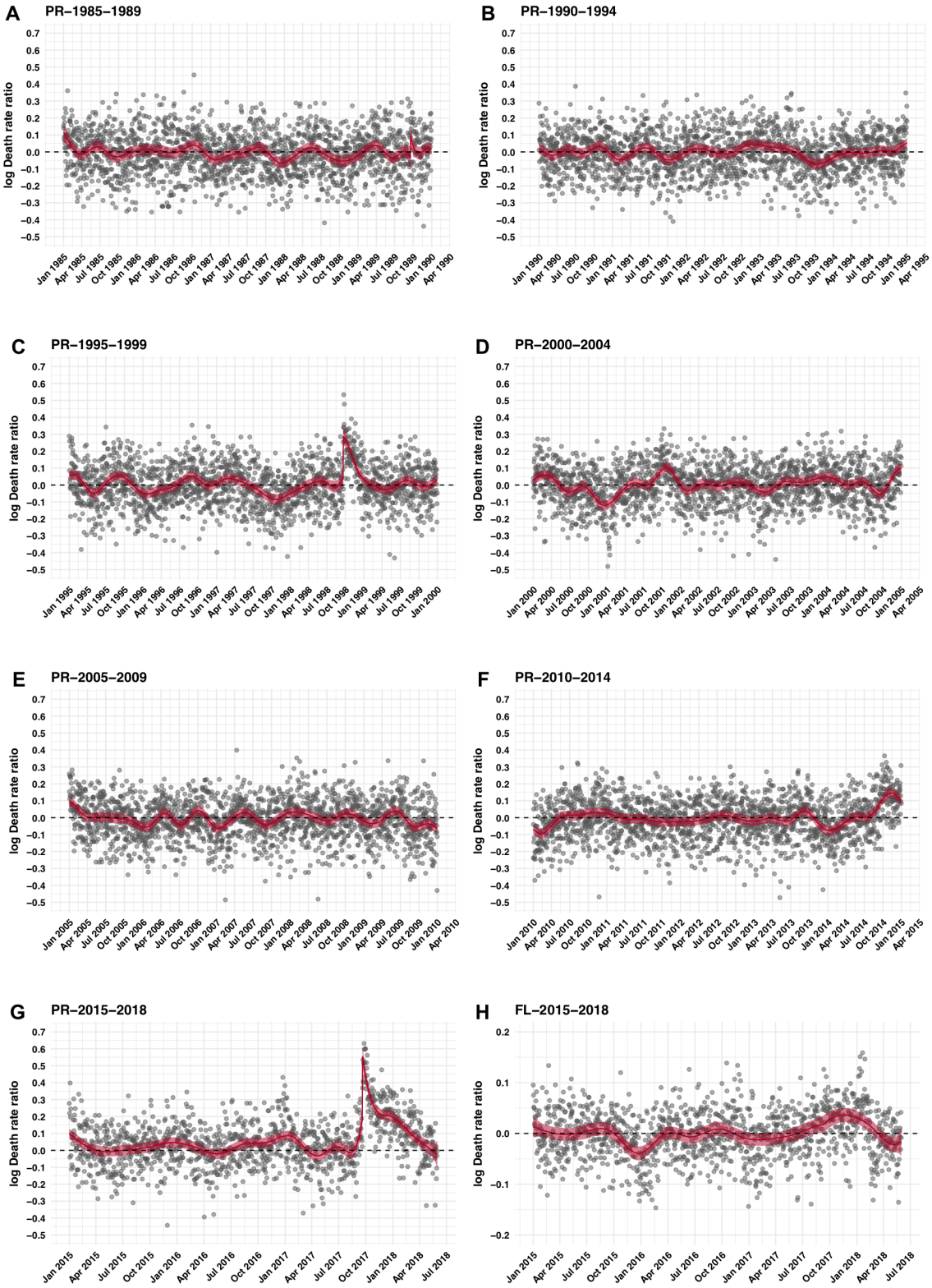

Figure 6: Model fit check: We compute the residuals after fitting the seasonal and yearly effects and plot these against time. The fitted $f(t_{i,j})$ and marginal 95% confidence interval are shown to capture trends. A) Puerto Rico 1985-1990 (includes Hugo). B) Puerto Rico 1990-1995. C) Puerto Rico 1995-2000 (includes Georges). D) Puerto Rico 2000-2005. E) Puerto Rico 2005-2010. F) Puerto Rico 2010-2015 (Includes Chigunguña epidemic). G) Puerto Rico 2015-2018 (includes María). H) Florida 2015-2018 (includes Irma)

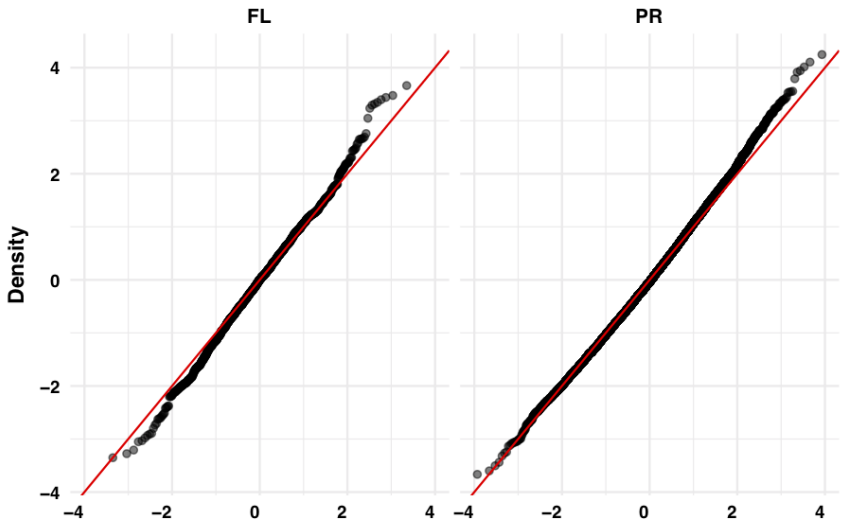

Figure 7: Poisson assumption check: we compute Pearson residuals and generate a normal qq-plot. We remove years determined to be outlier years. A) Puerto Rico. B) Florida

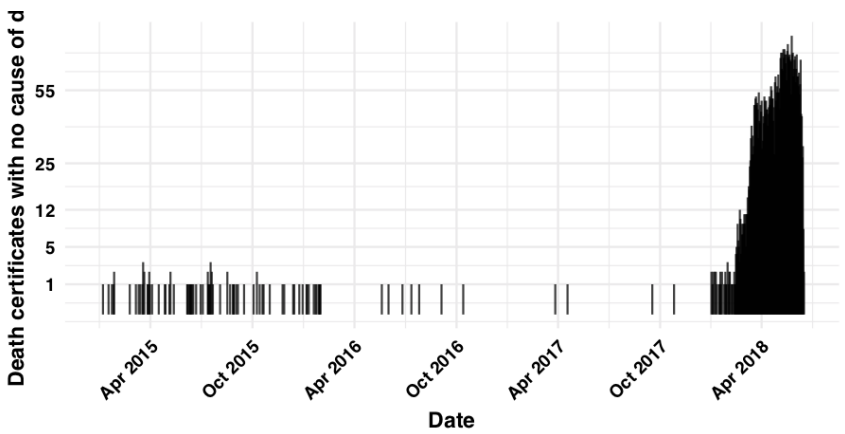

Figure 8: Number of individual records with no cause of death entered by day for the Puerto Rico individual 2015-2018 individual date

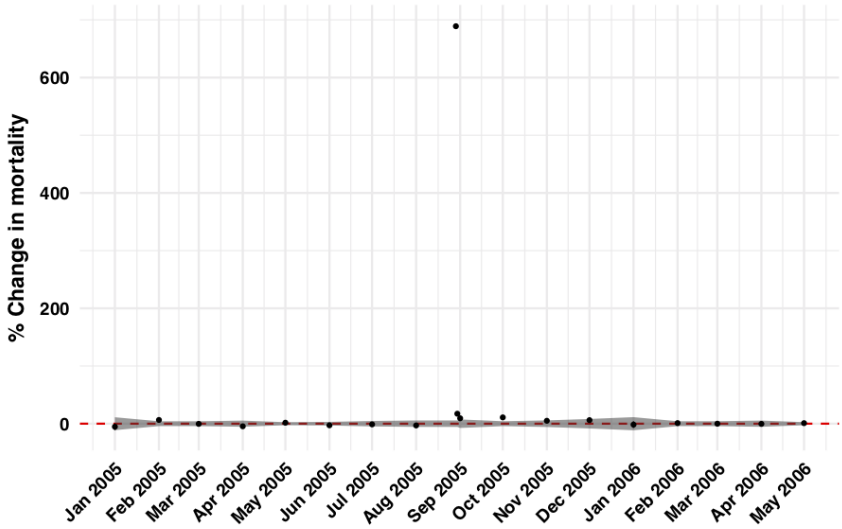

Figure 9: Estimated percent increases in death rates after hurricanes Katrina. Monthly estimates are provided except for August 29, 2005 and August 30-31, 2005 for which daily data was used to compute the number of deaths. The grey ribbon represents the range of variation (plus or minus two standard deviations) in death rate increases seen across non-hurricane

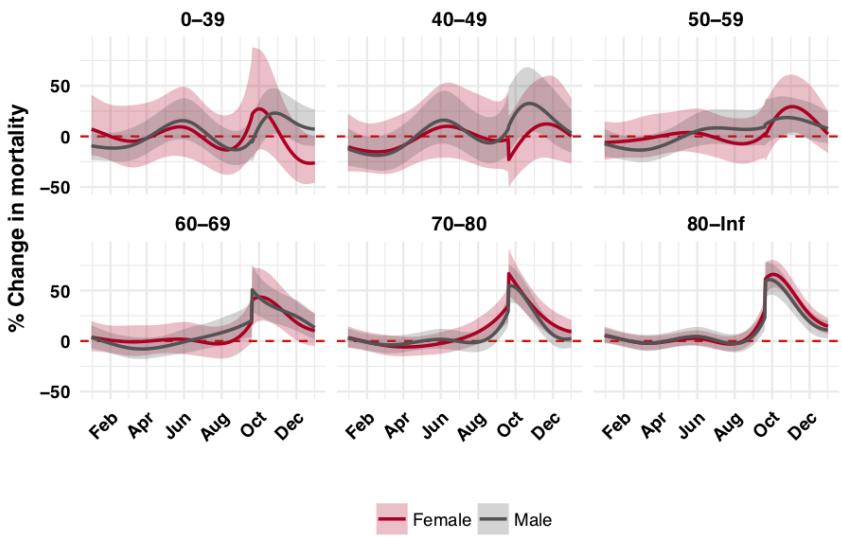

Figure 10: Estimated daily percent increase in death rate by age groups and gender

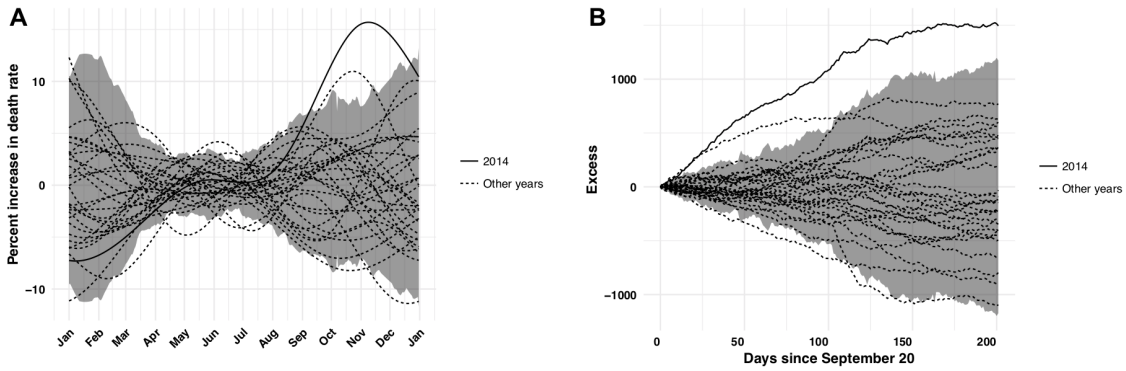

Figure 11: Natural non-hurricane related variation. For each day we estimated the standard deviation using median absolute deviation of f(ti,j). The dashed lines are the f(ti,j) and the grey ribbon is plus or minus two standard deviation.

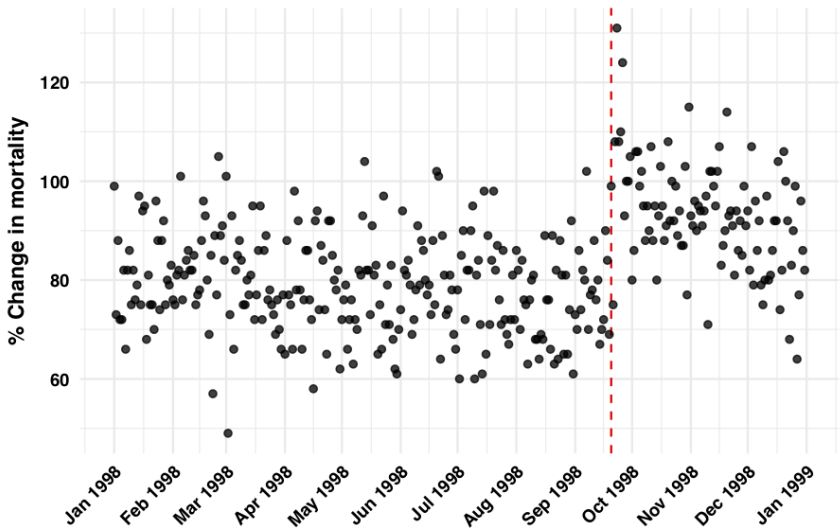

Figure 12: Estimated daily percent increase in death rate by age groups and gender