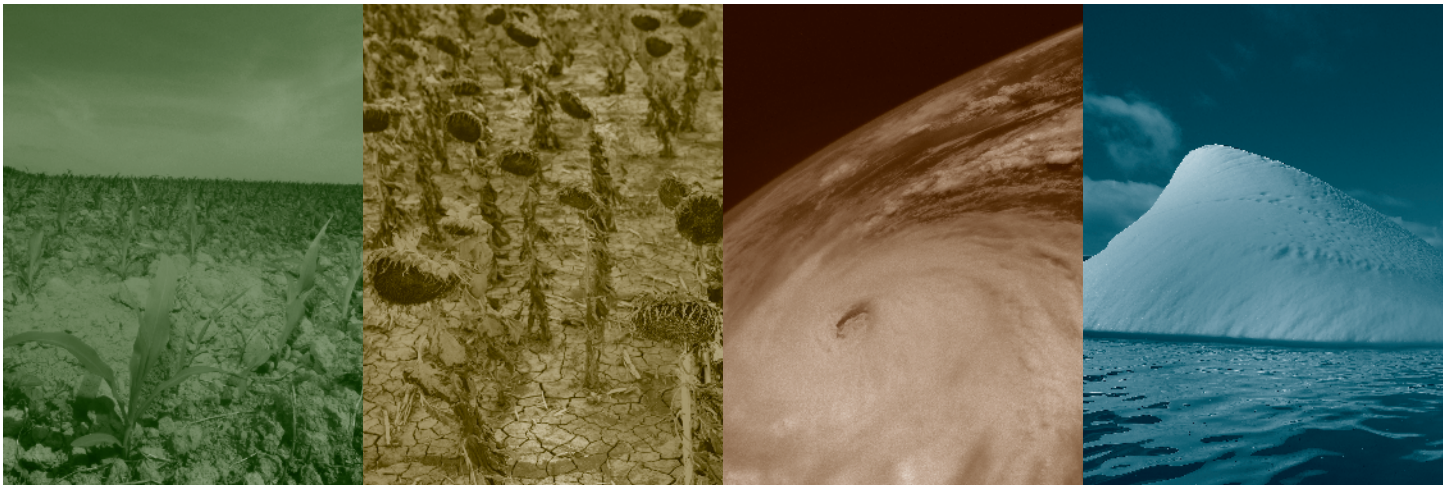
Technical Documentation on Exposure-Response Functions for Climate-Sensitive Health Outcomes



Climate and Health Technical Report Series

Climate and Health Program, Centers for Disease Control and Prevention

Kristina W. Kintziger,[1,2] Meredith A. Jagger,[1,3] Kathryn C. Conlon,[4] Kathleen F. Bush,[5] Brendon Haggerty,[6] Laurel Harduar Morano,[7] Kathryn Lane,[8] Matthew Roach,[9] Lauren Thie,[10] Christopher K. Uejio,[5,11] BRACE Methods Community of Practice (CoP)

[1]These authors contributed equally to this document.

[2]Florida Department of Health (former); Department of Public Health, University of Tennessee (current)

[3]Florida Department of Health (former); Acute and Communicable Disease Prevention Section, Public Health Division, Oregon Health Authority (current)

[4]Climate and Health Program, Division of Environmental Hazards and Health Effects (DEHHE), National Center for Environmental Health (NCEH), Centers for Disease Control and Prevention (CDC)

[5]Bureau of Public Health Statistics and Informatics, Division of Public Health Services, New Hampshire Department of Health and Human Services

[6]Oregon Health Authority (former); Multnomah County Health Department (current)

[7]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill

[8]Bureau of Environmental Surveillance and Policy, New York City Department of Health and Mental Hygiene

[9]Arizona Department of Health Services

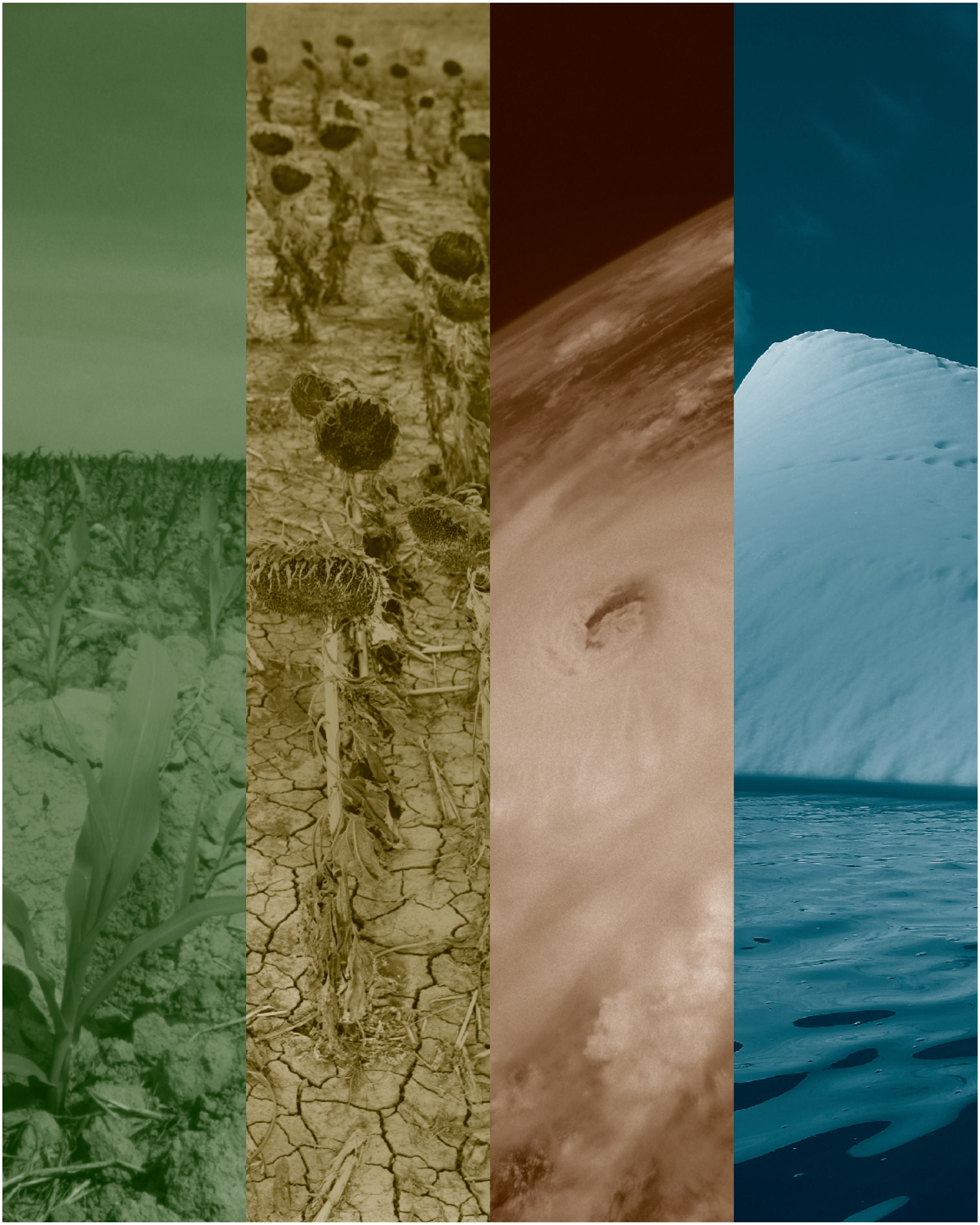
[10]Occupational and Environmental Epidemiology, Division of Public Health, North Carolina Department of Health and Human Services

[11]Department of Geography, Florida State University

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

National Center for Environmental Health
Division of Environmental Hazards and Health Effects

CS275363-A



1. Introduction

Climate change and climate variability influence human health in a variety of ways, and it may be important for the public health sector to consider these threats. Understanding the current and future burden of climate-sensitive diseases may benefit planning and response activities. As part of the Climate-Ready States and Cities Initiative (CRSCI), the Centers for Disease Control and Prevention (CDC) introduced the five-step iterative Building Resilience Against Climate Effects (BRACE) Framework (Figure 1) designed to help the public health sector plan for and adapt to climate-related hazards (http://www.cdc.gov/climateandhealth/BRACE.htm).

Part of BRACE Step 1 is focused on identifying potential climate impacts and associated health effects. Although this framework is considered to be an iterative process, each subsequent step builds on Step 1. Therefore, having a firm foundation, including an understanding of current risk and vulnerabilities, is key to subsequent steps, such as estimating current and projecting future disease burden.

In order to provide technical assistance and share best practices, the BRACE Methods Community of Practice (CoP) was developed to facilitate collaboration among CRSCI grantees and other partners interested in the connection between historic and future climate-sensitive health outcomes (e.g., heat-related illness [HRI], asthma, and vectorborne diseases). The purpose of this CoP is to explore methods to (1) quantify associations between climate-related environmental hazards and health outcomes, (2) incorporate these associations into projections of climate-sensitive health outcomes, and (3) develop best practices for public health agencies for assessing the future disease burden due to climate change.

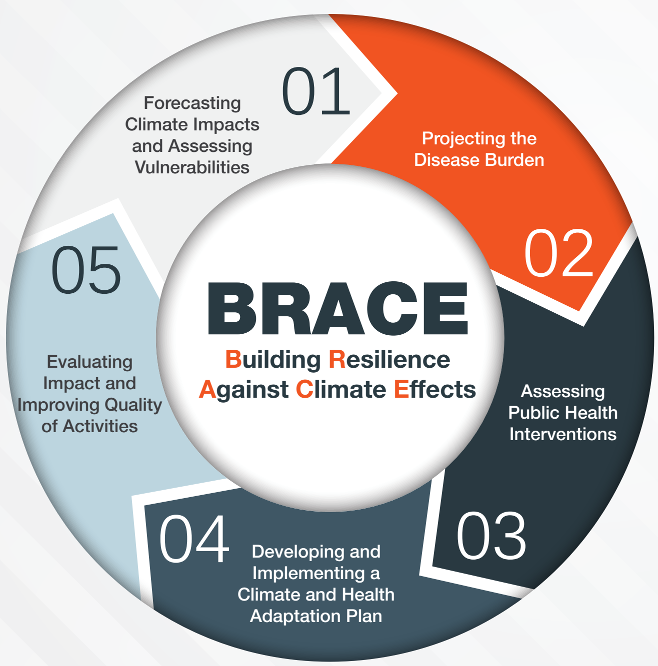


Figure 1. Building Resilience Against Climate Effects (BRACE) Framework

Recognizing the need for guidance for quantifying climate-related disease burden, the BRACE Methods CoP collaborated to (1) document available health, climate, and vulnerability data and (2) identify the epidemiological methods available to describe the potential associations between the climate-sensitive environmental hazard(s) and health outcome(s) (i.e., exposure-response functions). The purpose of this report is to provide resources to public health practitioners and agencies interested in applying the BRACE framework, including (1) examples of relevant health, weather, and climate data that can be used in health-effects studies, (2) provide resources on study designs commonly used to assess the relationships between climate hazards and health effects, and (3) provide case studies for quantifying the future disease burden due to climate change.

2. Health Data: Overview of Available National- and State-level Sources

Among the "10 Essential Public Health Services" identified by CDC, two services are focused on assessment through (1) "monitor[ing] health status to identify and solve community health problems" and (2) "diagnos[ing] and investigat[ing] health problems and health hazards in the community" (http://www.cdc.gov/nphpsp/essentialservices.html). These activities, undertaken at the local, state, and national levels, result in the collection of a vast amount of individual-level and aggregate health data. These data can be foundational for those working in the realm of climate and health (Figure 2). Epidemiologists working for a state or local health agency can generally access identifiable individual-level data for populations within their jurisdiction. External partners often require agency approval to access the de-identified data or aggregated statistics. In addition, other (non-health) agencies collect and share data that are critical to describing and studying population health. The utility of certain health datasets are contingent on the issue being addressed, including temporal and spatial scales. ***Thus, there is no universally accepted "best" health data source(s) for estimating exposure-response functions and conducting disease burden projections. All have strengths and limitations, and jurisdiction-specific priorities or constraints may direct choices.***

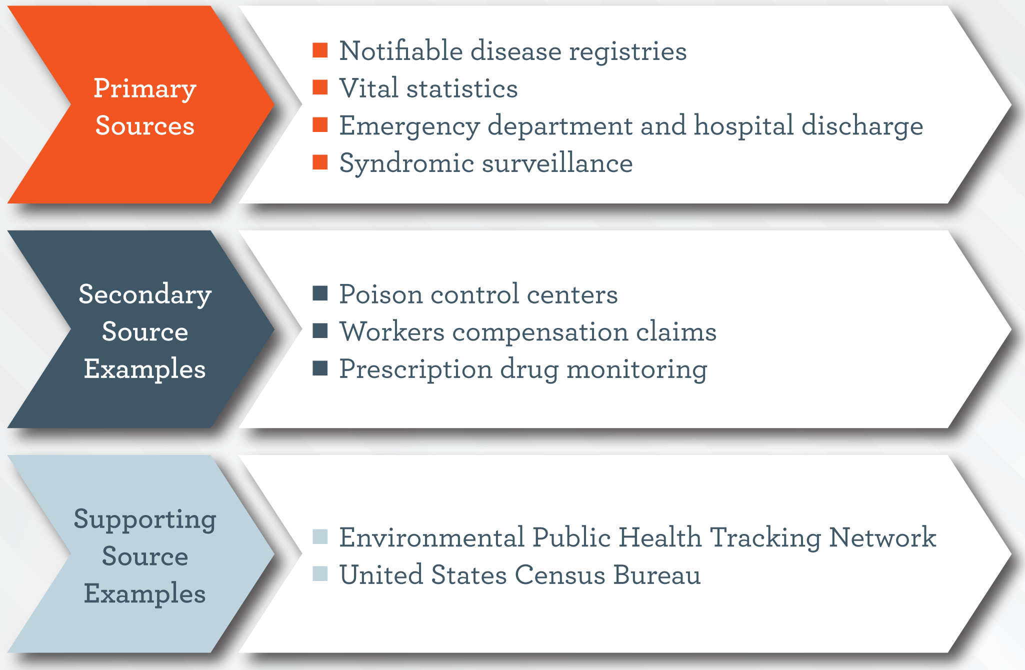


Figure 2. Examples of available health data sources for climate and health studies

2.1. Primary Health Data Sources

There are several available sources for health outcome data that may be used to examine the associations between climate and health. Please note that the data sources included in this document are those that have been most commonly used by BRACE grantees for developing exposure-response functions. Other data sources may be useful in other jurisdictions, but these are the primary sources currently available. Each source collects data for a variety of different purposes. To correctly and effectively use the health data, it can be helpful to have a good understanding of the data source and its associated strengths and limitations. Table 1 provides details for each of the commonly used primary data sources using one state, Florida, as an example to highlight the variation in the types of information available. However, available information will vary by jurisdiction.

DATA SOURCE	AVAILABLE DATES	GEOGRAPHIC RESOLUTION	TEMPORAL RESOLUTION	HEALTH VARIABLES	AVAILABILITY
Notifiable Disease Registry (Merlin)	2001–present	Residential address	Date of symptom onset, diagnosis, or report	Disease definition, exposure information, lab results, symptoms, outbreak associated, control measures, healthcare visits, travel history	As reported to health department
Vital Statistics	1917–present	Residential address	Date of death	Immediate and underlying causes of death (ICD-10 codes), contributing causes, tobacco use, manner of death	Up to 6-month lag
Emergency Department and Hospital Discharge	ED: 2005–present; Hospital: 1988-–present	Zip code	Date of visit or intake	Primary and secondary diagnoses (ICD-9/10 codes), primary and secondary procedures, external causes of injury (E- codes), payer category, billing charges	Up to 6-month lag
Syndromic Surveillance (ESSENCE)	2007–present	Zip code	Varies by source (e.g., date of intake)	Similar to other listed sources, but also includes chief complaint and discharge diagnoses	Varies by source (e.g., chief complaint and discharge data updated hourly to daily)

Table 1. Examples of commonly-used primary health data sources for climate and health studies used in Florida.

2.1.1. Notifiable Disease Registries

Local, state, and territorial health agencies have the responsibility of conducting surveillance on notifiable or reportable diseases and conditions. The list of specific reportable diseases and conditions varies by jurisdiction, but most are infectious or communicable. Health agencies maintain secure databases with individual case reports. In many jurisdictions, data are collected and entered primarily by county or local health department staff, although in smaller states, state health department staff may handle these responsibilities. A typical statewide database will allow real-time access for entering patient demographic and geographic information, case data (e.g., symptoms and possible exposures), laboratory results, healthcare visit information, extended case report form data, control measures, travel history, and outbreak information, where applicable. Most state and local reportable disease data are collected through passive surveillance.

Case data (i.e., nationally notifiable diseases) are voluntarily reported on a regular basis (e.g., weekly) to the CDC's National Notifiable Diseases Surveillance System (NNDSS; http://wwwn.cdc.gov/nndss/) or other reporting systems. NNDSS is a passive surveillance system, which collects de-identified data that have been stripped of certain protected health information that state and local jurisdictions collect for control and response activities. NNDSS uses a standard set of case definitions for reporting across all jurisdictions, though these definitions may change over time.

NNDSS and other CDC programs summarize and report these data in the *Morbidity and Mortality Weekly Report* (MMWR). Weekly U.S. morbidity tables from 1888 through 2013 have been digitized and made publicly available in an analysis-friendly format through the University of Pittsburgh's *Project Tycho*®
(http://www.tycho.pitt.edu/).

There are several strengths of notifiable disease registries that may be noted. Most are flexible and can be adapted to meet the unique needs for reporting of specific diseases. Robust systems have many years of data (i.e., at least 10, often more than 20), which provide important baseline information that may be used to assess disease rates and other trends over time. The limitations of notifiable disease registries include variation in training and expertise of staff (i.e., clinical knowledge, data entry and analysis experience), variable completeness, timeliness of case reporting, differing priorities for case follow-up, and differences in clinical and surveillance case definitions.

2.1.2. Vital Statistics

Vital statistics, including births, deaths, marriages, and divorces, are typically collected at the state or territorial level, often by the health agency. Here, we focus on the mortality data available within vital statistics systems. Many jurisdictions have been collecting mortality data since the early 20th Century, which contains the underlying (i.e., the disease or injury that initiated the train of events leading to death) and contributing causes (i.e., a significant condition that influences the course of the events leading to death, but not directly related to the disease or injury causing death) of death based on *International Classification of Diseases* (ICD) definitions (currently *Tenth Revision* or ICD-10 coding).

National mortality data are available through CDC's National Center for Health Statistics (NCHS), National Vital Statistics System (http://www.cdc.gov/nchs/deaths.htm). NCHS works with state and local jurisdictions to develop standard forms for the collection of vital statistics data to ensure uniform reporting. These data are made available through NCHS via public use data files and through CDC's online query system, WONDER (Wide-ranging Online Data for Epidemiologic Research: http://wonder.cdc.gov/). Fatal injury data is also made available through CDC's Web-based Injury Statistics Query and Reporting System (WISQARS: http://www.cdc.gov/injury/wisqars/index.html).

There are several strengths and weaknesses of vital statistics data that may be noted. The person completing death certificate information may vary by jurisdiction but may include the attending physician or one of their office representatives, a local medical examiner, or the funeral home director. There are quality control steps to ensure that the information reported is reliable and accurate. Limitations include the differing level of training or expertise for those completing death certificates, lag in availability, incomplete data, potentially miscoded data, changes to coding (i.e., conversion from ICD 9 to ICD 10 codes) and coding practices, and limited demographic or risk factor information.

2.1.3. Emergency Department and Hospital Discharge

At the state or territorial level, a variety of agencies collect data on emergency department (ED) visits and hospital discharges, including state health agencies, medical licensing agencies, or the agency that is responsible for Medicaid. These data sources contain a detailed record of each hospital inpatient, outpatient, and ED visit, and each record lists the primary and contributing diagnoses, patient demographics, and discharge information. Hospital discharge data may also contain information on primary and secondary procedures. Hospital inpatient and outpatient discharge data may be available for more years than ED, although many states have both all three available since the mid-2000s.

Limited hospital discharge and ED data are available at the national level. NCHS has several National Health Care Surveys (http://www.cdc.gov/nchs/dhcs.htm) that are conducted regularly (e.g., annually). These are designed to answer specific questions related to healthcare utilization and resource allocation and the quality of health care at the national level. The National Hospital Ambulatory Medical Care Survey (NHAMCS: http://www.cdc.gov/nchs/ahcd.htm) collects data from a nationally representative sample of visits to hospital EDs and outpatient departments to understand the utilization of hospital-based ambulatory care services. Data are currently available from 1992 through 2011. The National Hospital Discharge Survey (NHDS: http://www.cdc.gov/nchs/nhds.htm) collected data annually from 1965 through 2010 from a sample of inpatients who were discharged from non-Federal hospitals. NHDS has been replaced by the National Hospital Care Survey (NHCS: http://www.cdc.gov/nchs/nhcs.htm), which combines data from multiple NCHS surveys for more integrated analysis of hospital utilization trends.

ED and hospital discharge datasets have a number of strengths. These data provide fairly comprehensive statewide coverage, and many years of historical data are available. ED and hospital data provide epidemiologists the ability to study nonnotifiable diseases and injuries, and provide additional data to augment and evaluate notifiable disease information. Finally, these data provide overall and categorical healthcare charges that can be used to estimate the cost for weather-related health effects. Limitations to these data include the absence of data from federal and some state facilities, a lag (months to a year) in access to data due to internal reporting and validation processes, limited available identifiers, and questionable clinical accuracy as with any study relying solely on ICD-9 and ICD-10 codes.

2.1.4. Syndromic Surveillance

Syndromic surveillance is a tool that provides (near) real-time snapshots of population health data, which may function as either a primary or secondary health data source, depending on use. These systems allow for rapid identification, assessment, monitoring, and response to natural and man-made threats, and other events such as disease outbreaks. Public health practitioners, epidemiologists, emergency management officials, and environmental health experts play important roles in implementing and disseminating data collected from a syndromic surveillance system.

Syndromic surveillance data are collected from multiple sources, including: ED visits (triage notes, chief complaint data, ICD-9/10 codes), inpatient visits (chief complaint data, ICD-9/10 codes), ambulatory/outpatient, notifiable disease registries, vital statistics systems, and school/work absenteeism data. Data are normally collected electronically, through mobile data collection, or manually entered. Syndromic surveillance systems can vary but some of the most common platforms include: ESSENCE (Electronic Surveillance System for the Early Notification of Community-based Epidemics) (https://www.cdc.gov/mmwr/preview/mmwrhtml/su5301a30.htm) and RODS (Real-time Outbreak and Disease Surveillance) (https://www.rods.pitt.edu/site/). Health outcomes that are typically tracked by syndromic surveillance systems and relevant to climate include injuries, respiratory diseases, gastrointestinal illness, Lyme disease, West Nile virus, heat-related illness, cold-related illness, carbon monoxide poisoning, and asthma.

Limitations of the data relate to coverage and quality. Some syndromic surveillance systems rely on free-text search of discharge records, which can vary across providers and jurisdictions. Identification of common inclusion and exclusion criteria for spelling mistakes may be assessed. For example, with heat-related illness, common exclusion criteria are misspellings of “heart” and “head” (e.g., “heatache” or “heatbeat”). In addition, data are usually collected daily, but this varies. Geographic coverage of syndromic surveillance systems can also vary. These systems can be national, multi-state, state-specific, or smaller.

The use of syndromic surveillance systems to monitor the impacts of climate, extreme weather, and environmental exposures on population health is increasing. Evidence from a survey disseminated to public health syndromic surveillance staff around the country in July 2015 indicated that these systems are being used to track climate-sensitive health outcomes related to extreme heat, extreme cold, snow or ice, flooding, wildfire, hurricanes, tornadoes, poor air quality, and power outages.[1]

2.2. Secondary Health Data Sources

In addition to the primary data sources discussed in Section 2.1, there are a variety of secondary health data sources that may be used to augment or supplement health outcome data.

2.2.1. Poison Control Centers

The American Association of Poison Control Centers (AAPCC: http://www.aapcc.org) collects data from 57 poison control centers nationwide and maintains a nationwide surveillance database, the National Poison Data System (NPDS: http://www.aapcc.org/data-system/). Data are available from 1983 to the present and are reported annually in the NPDS Annual Report. Calls made to poison control centers may be related to areas of public health concern, such as notifiable diseases (e.g., food or water contamination events and associated illnesses) or environmental exposures (e.g., carbon monoxide poisoning). As such, these data are often used to enhance or support existing public health surveillance systems, and data may be made available to public health officials through a stand-alone system or incorporated into existing systems (e.g., syndromic surveillance systems). For example, carbon monoxide exposure calls may be useful for assessing climate-related events (e.g., exposure to generators post-hurricane or winter storm).

Often poison control centers are located on the campus of a major teaching hospital. Some centers provide services to multiple states. Patient exposures are assessed, managed, and coded by specialists in poison information, including pharmacists, nurses, physicians, or physician assistants trained and certified to operate the hotline. Data collected on each call include demographic and geographic information, date and site of exposure, reason for exposure, case management information, ingested substances, symptoms, and outcomes. Data are available in real-time, 24 hours per day and seven days per week. Of note, while certified centers have follow-up protocols to ensure accuracy and completeness of data, most of the data are self-reported and missing information can be a problem.

2.2.2. Workers Compensation Claims

Many states require reporting of workers compensation claims to a state agency. These claims may be of use to climate and health programs for surveillance related to climate hazards, especially claims of first responders, such as wildland firefighters, and vulnerable outdoor workers. Reporting requirements and administering agencies vary by state, but basic information is available from the Federal Department of Labor (http://www.dol.gov/owcp/dfec/regs/compliance/wc.htm) and the Occupational Safety and Health Administration (OSHA) (https://www.osha.gov/workers/index.html). Many states also have an Occupational Health Surveillance Program that works in collaboration with the National Institute for Occupational Safety and Health (NIOSH). An important consideration related to reporting requirements is whether all claims are reported, just those that are accepted (not denied), or just those that are disabling. A reporting requirement that is limited to accepted disabling claims would exclude minor injuries that may be of interest to investigators. In such cases, data use agreements can be arranged between health departments and workers compensation insurers.

2.2.3. Prescription Drug Monitoring Programs

Almost all states have enacted a prescription drug monitoring program. Typically, these programs are created by state legislatures to track the use of scheduled substances with the goal of reducing drug abuse. These programs are often administered by boards of pharmacy or health departments, but sometimes by other agencies such as a law enforcement agency. Many states provide de-identified data for research purposes. Some states collect data on schedules II-V, while others collect data on a more limited set of substances, schedules II-IV. Drug schedules are a classification system that categorize substances based on medical use and dependency potential, with lower numbers indicating greater potential for abuse and dependency (http://www.dea.gov/druginfo/ds.shtml). Additionally, some states choose to track additional substances such as pseudoephedrine, which is sometimes used in the manufacturing of methamphetamine. Prescriptions or refills could inform an exposure-response association for some climate-related hazards. For example, researchers in British Columbia have associated wildfire smoke with albuterol dispensations for asthma patients.[2] The Prescription Drug Monitoring Program Training and Technical Assistance Center at Brandeis University maintains information about programs in each state (http://www.pdmpassist.org/).

2.3. Supporting Data Sources

Unlike the primary and secondary health data sources above, which can include record-level information on health events (e.g., deaths, ED visits, poison center calls), supporting health data sources typically include only population-level aggregate information.

2.3.1 Environmental Public Health Tracking

The Environmental Public Health Tracking (EPHT) Program (http://ephtracking.cdc.gov/), implemented by the Centers for Disease Control and Prevention (CDC) in 2002, is focused on collecting, integrating, analyzing, and interpreting data for environmental hazards, human exposure, and the health effects potentially related to exposure. The ultimate goal of this ongoing, systematic collection of data is to plan, implement, and evaluate public health action. The National EPHT Program, also known as the National Tracking Program, includes grantees from 25 states and one city, as well as academic partners and Peer Fellowship Program participants. Each grantee maintains a State (or City) Tracking Portal (http://ephtracking.cdc.gov/showStateTracking.action) to provide state-specific resources to local partners, and reports annual data to the CDC for publication on the National Tracking Portal. Tracking portals include both health outcome data (described in this section), and hazard and environmental exposure data (described in Section 3.2).

As of 2015, health indicators potentially related to climate and weather include asthma, heat-related illness, and carbon monoxide poisoning counts and rates aggregated at varying spatial and temporal resolutions. A variety of other health conditions potentially related to the environment are also available. Sources of health data for the Tracking Portals include birth defects and cancer registries, childhood lead poisoning prevention programs, emergency department and hospital inpatient data, and vital statistics. Most of these are nationally-required Tracking indicators available for all grantees. Additionally, jurisdictions may provide state-specific indicators on their portals that are of interest locally (e.g., built environment, harmful algal blooms). New recommendations for nationally consistent data measures (NCDMs) related to climate and weather are under consideration for the National and State Tracking Portals, including indicators for temperature and Lyme disease. Additionally, using case definitions from EPHT NCDMs can be helpful in querying new data (ex. heat illness cases from hospital discharge datasets) for climate-related analysis.

A major advantage of the EPHT network is that common case definitions have been developed and validated, so data are comparable across jurisdictions. Strengths and limitations of each indicator are detailed in the metadata. A limitation is that data may only be available at the county level to protect patient confidentiality, and are released in pre-defined formats, for instance age groupings, that may not meet the specific requirements of researchers. The State and National Tracking Portals also include data from the U.S. Census Bureau related to housing, population, and poverty.

2.3.2. U.S. Census Bureau Population Data

There are several population-based surveys conducted by the U.S. Census Bureau, part of the U.S. Department of Commerce, on a regular basis. Most of these data are free and available for download and use from the website (http://www.census.gov or http://factfinder2.census.gov). County, ZIP code-, and tract-level data are available for population information. However, ZIP code-level data are based on ZIP Code Tabulation Areas (ZCTAs), which are generalized representations of the ZIP code service areas used by the Census Bureau.

2.3.2.1. Decennial Census

The most well-known of these surveys is the decennial census, which is conducted every 10 years nationwide and is constitutionally mandated (Article 1, Section 2 of the U.S. Constitution). A number of previous enumerations included a short form survey, which only collected basic demographic information and a long form survey (sent to around one in every six U.S. households) which collected more in-depth demographic and housing information. In 2010, a single form containing four core household-level questions and five individual-level questions per household member was used and the long form survey was discontinued. The most recent decennial census data are used to estimate population projections for future years.

2.3.2.2 American Community Survey (ACS)

This survey is conducted annually throughout the U.S. Since the long form was retired from the decennial census after 2000, the ACS now collects the same information from a sample of approximately 3.5 million residents. These data include age, gender, race, ethnicity, family and relationships, income, benefits, health insurance, education, veteran status, disabilities, occupation, and cost of living.

3. Climate Data: Overview of Available National- and State-level Sources

For epidemiologists developing exposure-response functions for climate-sensitive health outcomes, high quality climate data may be very valuable. However, there are technological challenges that those working in public health, especially at state and local health departments, may encounter when working with these data. While building collaborative relationships with local meteorology or climatology experts can be very helpful, there is a basic level of understanding that may be useful for evaluating available data sources. For example, historical weather station data must be aggregated spatially and temporally to match the resolution of health data. Modeled climate datasets can be quite large and usually come in formats (e.g., netCDF) that are unfamiliar to epidemiologists, and require the use of software packages and languages to which many public health practitioners may be unaccustomed. A discussion on technical challenges related to working with climate data is beyond the scope of this report. However, connecting with technical experts to discuss problems and solutions is within the scope of the BRACE Methods CoP. This section contains some of the publically and freely available data sources and resources which have been used by CRSCI grantees in quantifying exposure-response functions for climate-sensitive health outcomes.

3.1 National Oceanic Atmospheric Association (NOAA)/ National Weather Service (NWS) Observed Ground-Level Weather Data

We briefly review the technical definitions of common weather metrics used in human health studies. Changes in the atmosphere's energy or moisture may directly or indirectly impact human health. Surface air temperature is the kinetic energy of air molecules at a height of 2 m above the ground. Temperatures at this height are relatively stable compared to the temperature closer to the surface. Precipitation is the condensation of water vapor that falls from the atmosphere via gravity such as rain droplets, hail, graupel, or snow. Multiple metrics attempt to measure the amount of moisture in the air which may influence human thermal comfort.[3] Relative humidity is the amount of water vapor in the air divided by the atmosphere's capacity to hold water vapor at a given temperature. Thus, relative humidity depends on both moisture and temperature, while other metrics such as dew point or specific humidity only measure water vapor. Wind speed is the ratio of the distance covered by the air to the time taken to cover it.[4] Intuitively, wind direction is the direction from which the wind is blowing.

A comprehensive review of observed, modeled, and remotely sensed weather and climate data sources is beyond the scope of this report. Table 2 summarizes key characteristics from observed ground-level U.S. weather information. The data sources are freely available for government, academic, and/or public research purposes. Notably, observed ground-level information is compiled by a wide range of institutions. NOAA's National Centers for Environmental Information (NCEI) hosts the largest repository of information. However, these data sources can be augmented by information collected by other agencies.

Selecting the most appropriate dataset depends on the study question of interest. There are implicit trade-offs between the quality, frequency, historical length, and variety of weather conditions collected by each dataset. For example, the primary purpose of the high-quality United States Historical Climatology Network (USHCN) (https://www.ncdc.noaa.gov/oa/climate/research/ushcn/) is to detect regional climate change. Quality control procedures verify that stations have not moved locations, remove outliers, and adjust for potential discontinuities (e.g., changing the type of thermometer). The network selected stations outside of cities that may be less influenced by urban growth and land cover change. However, a public health practitioner may be more interested in conditions inside of cities where most of the population resides.

A hypothetical aeroallergen and respiratory health study provides another illustration of the trade-offs between data sources. Temperature and precipitation, as well as relatively infrequently collected humidity and wind speed and direction, may influence aeroallergen exposure. NWS employees or certified observers collect high-quality, frequent, and comprehensive information at "first order" stations contained in the Quality Controlled Local Climatological Data. First order stations can be found at airports and other locations (e.g., universities, public parks). However, there are relatively few first order stations. Thus, the study may blend information from both the Quality Controlled Local Climatological Data (https://www.ncdc.noaa.gov/qclcd/QCLCD?prior=N) and the more ubiquitous Global Historical Climatology Network (https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-ghcn).

Some NOAA NCEI datasets compile information from multiple observing networks. For instance, the Global Historical Climatology Network Daily (GHCN-D) database (ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/) contains observations from approximately 30 different data sources. In the U.S., this includes the Cooperative, First Order, ASOS Summary of the Day (https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/automated-surface-observing-system-asos), the Climate Reference Network, and CoCoRaHS (Community Collaborative Rain, Hail, and Snow Network), which is a network of volunteer weather observers throughout the U.S. and Canada. While quality assurance is applied to the GHCN-D observations, they are not homogenized to a uniform standard. Intuitively, the component datasets are subject to the limitations and constraints of the original data set. Thus, public health practitioners can read the metadata and associated publications to learn about the limitations of each dataset.

DATA SOURCE	AVAILABLE DATES	TEMPORAL RESOLUTION	WEATHER VARIABLES	REFERENCES
United States Historical Climatology Network (USHCN)	~1880-present	Daily, Monthly	Daily (temperature, precipitation, snowfall, snow depth), monthly (temperature, precipitation)	5–9
Global Historical Climatology Network (GHCN)	Variable	Daily, Monthly	Total precipitation, temperature. Sometimes: snowfall and snow depth	10–12
National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI) Quality Controlled Local Climatological Data	1996–present	Hourly, Daily, Monthly	Temperature, precipitation, relative humidity, wind speed and direction, cloud cover	13, 14
NOAA NCEI Integrated Surface Data	~1900–present	Hourly	Temperature, precipitation, relative humidity, wind speed and direction, wind gust, cloud cover, pressure, weather type, visibility	15, 16
NOAA National Weather Service (NWS) Cooperative Data	1948–present, oldest - 1850's	Daily, Monthly, Annual	Temperature, precipitation. Sometimes: snow depth, snow water equivalent	17–19
NOAA NCEI Global Surface Summary of the Day	1929–present	Daily	Temperature, dew point, pressure, wind speed, precipitation, snow depth	20
United States Geological Survey (USGS) Surface Water Data	1889–present	≤Daily, Monthly, Annual	Water level/flow (e.g., streamflow, gauge height). Sometimes: water quality (e.g., ammonia, dissolved oxygen, nitrate, turbidity), meteorology (e.g. precipitation, temperature)	21, 22
USGS Instantaneous Data Archive	1984–present	≤Hourly	Discharge	21, 22
National Trends Network (NTN)/National Atmospheric Deposition Program (NADP)	1978–present	Daily, Seasonal	Precipitation, free acidity (H^+ as pH), conductance, calcium (Ca^{2+}), magnesium (Mg^{2+}), sodium (Na^+), potassium (K^+), sulfate (SO_4^{2-}), nitrate (NO_3^+), chloride (Cl^-), and ammonium (NH_4^+), total mercury (Hg)	23
United States Department of Transportation CLARUS	2009–present	Hourly	Temperature, precipitation, relative humidity, wind speed and direction, surface covered with ice	24

Table 2. Example of observed, ground level weather information collected by U.S. federal agencies.

3.2. Environmental Public Health Tracking—Climate Data

The National and State Tracking Portals (http://ephtracking.cdc.gov) were described in detail in Section 2.3.1. Here, we describe the hazard data available at the national and state levels related to climate and weather. In general, environmental exposure and hazard data include air quality, weather, and drinking water.

As of 2015, there were several indicators available on the National Tracking Portal related to temperature. Temperature, heat index, and the number of extreme heat days are available nationwide at the county level for both historical and future time periods. Historical data were obtained from the North American Land Data Assimilation System (NLDAS) (http://ldas.gsfc.nasa.gov/nldas/) for May–September and 2000–present, and include temperature and heat index (both in degrees Fahrenheit) and extreme heat days. For extreme heat days, the absolute thresholds include 95°F, 100°F, and 105°F, while relative thresholds include 90th, 95th, and 98th percentiles determined using both temperature and heat index, calculated based on the months of May through September. For extreme heat events, the absolute thresholds include 90°F, 95°F, 100°F, and 105°F, while relative thresholds include only the 90th, 95th, and 98th percentiles for 2 or more consecutive days and 3 or more consecutive days. Historical daily estimates of maximum temperature and heat index for the summer months (May–September) are also provided using NLDAS data from 1979 to the present.

Temperature projections were obtained from statistically downscaled global circulation model data, specifically the Statistical Asynchronous Regional Regression Daily Downscaled Climate Projections: 1/8 degree-CONUS Daily Downscaled Climate Projections (https://cida.usgs.gov/thredds/catalog.html?dataset=cida.usgs.gov/thredds/dcp/conus_pr) by Dr. Katharine Hayhoe. Variables include projected number of future extreme heat days and nights, identified based on both absolute and relative thresholds for both the A2 and B1 emissions scenarios. For extreme heat days, the absolute thresholds provided include 90°F, 95°F, and 100°F, while relative thresholds include 90th, 98th, and 99th percentiles. For extreme heat nights, the absolute thresholds include 65°F, 75°F, and 85°F, while relative thresholds include only the 98th percentile. Projections are available for the years 2020 through 2099.

As of 2017, additional exposure data related to climate are proposed to be included on the National Portal, including additional measures of heat exposure, as well as wildland fires and extreme weather events.

3.3. PRISM Climate Data

The PRISM Climate Group at Oregon State University publishes gridded datasets of precipitation and temperature for the U.S. These data are some of the most widely used spatial climate datasets and are the official spatial climate datasets of the U.S. Department of Agriculture. PRISM assumes that elevation is the most important factor in the distribution of climate variables, combining station observations with elevation and physiographic variables to produce high-resolution datasets.

This overview is intended to provide basic familiarity for public health professionals. Those interested in using the data can refer to the extensive documentation provided on the PRISM website (www.prism.oregonstate.edu). PRISM includes the variables listed below. Availability depends on the dataset: precipitation, temperature (mean, minimum, and maximum), mean dew point temperature, vapor pressure, vapor pressure deficit (minimum and maximum), and elevation. PRISM data are available in 4 km or 800 m grid cell sizes. Datasets at the 4 km resolution are available free of charge for all time scales, whereas monthly or daily 800 m data must be ordered from the Northwest Alliance for Computational Science and Engineering.

The PRISM Explorer interface allows users to retrieve data for county centroids (geographic mid-points) or user-entered latitude and longitude coordinates. With the exception of statewide precipitation maps, no summaries are available for political boundaries such as states, counties, or census tracts, but the data can be processed in GIS or other software to create such summaries.

PRISM data are available on time scales ranging from 30-year normal to daily values. Availability is summarized in Table 3 below.

TIME PERIOD	AVAILABILITY
1895–1980	Annual and monthly
1981–present	Annual, monthly, and daily
1981–2010	Annual and monthly 30-year climate normals
Prior 6 months	Monthly, daily provisional results

Table 3. Data available from PRISM.

The PRISM Climate Group at Oregon State University periodically updates the data, web interfaces, and supporting documentation. They offer most data products freely, although they may request a fee for more complex data needs. As noted extensively in the available documentation, PRISM data less than six months old are provisional and likely to change. The dataset documentation also states, “These datasets are not static entities, but are in a constant state of change. New networks are being added periodically to some datasets. Even those designed for long-term consistency experience changes due to improvements in data handling and quality control procedures.”

PRISM data are “modeled” in the sense that they are produced from a statistical model that estimates climate variables spatially between observations. This is not to be confused with “modeled” in the sense of future projections. Therefore, they may be useful in establishing exposure-response relationships or informing surveillance activities, but it may be helpful to apply caution if comparing to climate projections. Readers hoping to compare PRISM data to climate projections are advised to consult a climatologist.

PRISM is a very versatile data source and is excellent for describing recent climate and establishing associations between climate variables and health variables. Among PRISM’s limitations is that its reference period for climate normals is 1981–2010 and cannot be altered. This time period is useful for describing the recent climate, but may not align with other reference periods that are intended to capture conditions more typical of the 20th century. While many of the products are easily accessible for use in a variety of software applications, some knowledge of GIS may be useful for summarizing data for larger geographies or processing large numbers of records.

There are multiple ways to access PRISM data online. Among these are menu-driven user-friendly interfaces that allow downloads in multiple formats depending on the dataset desired (.asc, .bil, .csv, and .png images). Metadata are readily available in xml format. Bulk downloads are available via FTP (file transfer protocol) or web service.

3.4. Coastal Resources

3.4.1. Comprehensive Surge Database (SURGEDAT)

SURGEDAT is a comprehensive database for storm surge data for the U.S., developed and maintained by the Department of Geography and Anthropology at Louisiana State University (LSU) (http://surge.srcc.lsu.edu). It contains several unique datasets of interest, including U.S. Gulf of Mexico Peak Surge, East Coast Surge, and Super SURGEDAT.

3.4.1.1. U.S. Gulf of Mexico Peak Surge Database

Until recently, SURGEDAT was focused on the U.S. Gulf of Mexico. For the Gulf Coast, SURGEDAT has collected data on location and peak surge height for all surge events since 1880 with at least a 1.22 m (4 feet) storm surge. More recently, the database has begun to include information on surge events that were less than 1.22 m high. Information includes location and coordinates, storm system, and peak surge height, as well as peak surge-tide height. A variety of data sources were used to compile SURGEDAT including federal government agencies, academic publications, and newspaper and periodical articles.[25] Of note, for many systems, only the peak surge event and location are included and not the full extent of the storm surge for each system. Therefore, other areas may have experienced significant storm surge during an event but are not included in the database, particularly for older storm systems. Other key limitations are the following: tide heights are based on both scientific and anecdotal information; and the database is unable to differentiate between datums, references of mean sea level, and storm tide versus storm surge, in some instances. This database may be downloaded as a CSV file from the website (http://surge.srcc.lsu.edu/files/gompeaksurgedb.csv).

3.4.1.2. East Coast Surge Data

More recently, SURGEDAT has been expanded to include storms affecting the U.S. East Coast. According to the website, the East Coast database has data available for over 75 storm surge events and historic surge envelopes (i.e., area affected by surge) for over 40 storms. The website states a completion date for this database of 2013. The data are not currently available on the website, but researchers at LSU can be contacted about this database through the Contact Us link (http://surge.srcc.lsu.edu/contact.html).

3.4.1.3. Super SURGEDAT

This database contains historic surge envelopes, surge inundation maps, and a return frequency analysis. Historic surge envelopes use "historic surge observations to draw a spline-interpolated high-water envelope along an entire coastal region." This allows for storm surge estimates even in areas with sparse data. The surge inundation maps visually depict the maximum storm surge or storm tide height associated with given storms along entire coastal areas. The return frequency analysis is an experimental tool that "estimates the return period of storm surge heights in specific locations." For more information on these items in Super SURGEDAT, contact researchers at LSU (http://surge.srcc.lsu.edu/contact.html).

3.4.2. Surging Seas: Sea Level Rise Analysis by Climate Central

Surging Seas is a program by Climate Central that focuses on coastal flood hazards and rising seas, and seeks to provide local level surge risk analysis and tools to all of the U.S. coastal states (http://sealevel.climatecentral.org/). Surging Seas, itself, has two main tools available for many coastal states including the Surging Seas Risk Finder and the Submergence Risk Map. The Risk Finder provides information on population, infrastructure, and assets that may be at risk for coastal flooding hazards, and allows the user to compare risk across areas and determine the likelihood of these risks in the future. The Risk Map "shows areas vulnerable to flooding from combined sea level risk, storm surge, and tides, or to permanent submergence by long-term sea level rise."

In addition to the specific tools developed by Surging Seas, the website also provides a data matrix (http://sealevel.climatecentral.org/matrix/) that compares and provides links to other coastal resources by state. For example, most states also have links for available tools that cover their coastline from NOAA's Office for Coastal Management (Sea Level Rise and Coastal Flooding Impacts Viewer) and The Nature Conservancy (Coastal Resilience Mapping Tool), discussed in depth below. In addition, some states have additional coastal resources relevant to the hazards of coastal flooding and sea level rise. Links and descriptions to these tools have also been provided. For example, the California matrix (http://sealevel.climatecentral.org/matrix/CA.html) has links to tools at the Pacific Institute, Cal-Adapt, and "Our Coast, Our Future." State-specific resources are not described in this document, and can be located within the matrix for the specific state of interest.

3.4.3. NOAA's Office for Coastal Management, Sea Level Rise and Coastal Flooding Impacts Viewer

The Sea Level Rise and Coastal Flooding Impacts Viewer (http://coast.noaa.gov/slr/) is part of NOAA's Digital Coast program (http://coast.noaa.gov/digitalcoast/), which provides a wide variety of data, tools, and training to the coastal management community. The Viewer is intended to provide coastal managers and scientists with maps of the potential coastal impacts of sea level rise and related information for community officials. The maps provide the ability to simulate different sea level rise scenarios, anywhere from one to six feet over average high tides. Users are also able to determine how various landmarks may be affected at the different scenarios.

There is also information on marsh impacts, including modeling the potential marsh migration that may occur due to sea level rise, nuisance flood frequency, and even socioeconomic and vulnerability data. Data are currently not available for Alaska and Louisiana due to issues in the accuracy of available elevation data and the complexity of the coastlines of these states.

3.4.4. The Nature Conservancy, Coastal Resilience Mapping Tool

The Nature Conservancy provides a tool through Coastal Resilience (http://coastalresilience.org) that helps communities understand their vulnerability to coastal hazards. Coastal Resilience is intended to support communities and practitioners around the world in applying planning innovations to reduce community risk to coastal hazards by implementing mitigation and adaptation strategies, including working within natural ecosystems to find solutions (e.g., planting mangroves, creating coral reefs). Some of the resources used by the Coastal Resilience mapping tool (http://maps.coastalresilience.org/network/) include: Global Platform on Risk Reduction, World Risk Report, and Conservation Atlas. The mapping tool includes information such as natural defense projects in a variety of habitats and their role in coastal protection, coastal flood hazards, and sea level rise (zero to six feet), as well a map layers for habitats, hazards (e.g., storm hazards, waves and sea level, and socioeconomic data), and ecoregions.

3.5. State-Specific Resources

States and local jurisdictions can also take advantage of locally-specific environmental exposure, weather, and climate data for use in health outcome and projection analyses. Consortiums of researchers and climatologists may make regionally-specific climate projections available. An advantage of these data sources is that they may have a higher temporal and spatial resolution. At the same time, as with any data source, there are limitations for each, and, for data that are not publicly available, use must be negotiated with the appropriate data stewards. The purpose of the examples below is not to provide an exhaustive accounting of state-level data sources, but rather to highlight the range of possibilities for public health practitioners who would like to access similar state-level resources. We suggest speaking with local environmental health practitioners or climatologists to determine available jurisdiction-specific resources.

3.5.1. New York City Panel on Climate Change (NPCC)

The NPCC is a group of scientists, climatologists, and experts that specialize in risk management, social science, and climate adaptation who are charged with the development of downscaled projections of climate and assessment of potential impacts for New York City. First convened in 2008, the panel was mandated to meet regularly to update climate predictions and assessments for the city in 2012 by a NYC City Council local law. The panel also advises the Mayor's Office of Sustainability and the Mayor's Office of Recovery and Resiliency (ORR) on communication about climate projections to NYC residents. The panel publishes regular summary reports about the projections, which include average quantitative changes for variables such as temperature and precipitation qualitative assessments of climate variables such as coastal storms.[26] The panel regularly updates projections and incorporates new climate science into its assessments. Limitations are detailed in their public reports, including uncertainly about future greenhouse gas emission scenarios, potential for downscaling techniques to fail to capture some types of climate processes, and uncertainty about natural variation, among others. Researchers who would like to work with finer scale data may request them directly from the panel.

3.5.2. North Carolina Environment and Climate Observing Network (ECONet)

The North Carolina Environment and Climate Observing Network (ECONet) is a research quality meteorological observation network of 40 stations located across the state of North Carolina hosted by the State Climate Office of North Carolina (http://climate.ncsu.edu/econet). The NC ECONet provides weather data for the complex and ever changing climate of North Carolina. Weather stations in the ECONet are stand-alone towers that measure 10 meters tall with meteorological sensors at 2 m, 6 m, and 10 m with additional soil sensors up to 20 cm below the surface. Data are recorded on 1-minute time scales and collected at various intervals for processing, and are then used by the NWS and other collaborators.

3.5.3. North Carolina Heat-Health Vulnerability Tool

North Carolina has county-level temperature (and soon to be added heat index) and heat-related ED data available in the Heat-Health Vulnerability Tool (http://sercc.com/hhvt). Further, the tool models expected heat-related emergency department visits based on forecasted temperatures. This tool was funded by the Carolinas Integrated Sciences Assessment, the North Carolina BRACE grant, NC DETECTS (North Carolina Disease Event Tracking and Epidemiologic Collection Tool—North Carolina's syndromic surveillance system), the North Carolina State Climate Office, and the Southeast Regional Climate Center.

3.5.4. Florida Automated Weather Network (FAWN)

FAWN was established and is maintained by the University of Florida's Institute of Food and Agricultural Sciences (IFAS) (http://fawn.ifas.ufl.edu). FAWN received state legislative fund appropriations in 1997 and was initiated at the University of Florida Institute of Food and Agricultural Sciences into an existing county Cooperative Extension Service network. FAWN has been collecting data since the mid-1990s to present on various indicators related to soil and weather, and initially included 11 sites but has since expanded to include 40 sites throughout rural Florida. FAWN collects sub-hourly to annual weather data for stations located in agricultural areas around the state. Temperature is measured in degrees Celsius, 2 meters from the ground, and relative humidity is also measured at each of these stations. This network exists to supplement weather data observed by the NWS observation stations. These data are not reported to NCEI, and constitute a separate system that fills in important weather-related gaps around the less densely populated areas of the state.

4. Overview of Available Study Designs and Statistical Analyses in a Climate and Health Context

There are many epidemiologic study designs available for examining exposure-response functions. However, only a few are appropriate for and commonly used in climate and health studies. As with any epidemiologic investigation, there are several key considerations when choosing a study design: the type of health data (e.g., individual records or aggregate counts), the temporal and spatial scale of the data, and the type of exposure (e.g., acute, transient, long-term). While it is outside the scope of this project to provide the detailed statistical theory associated with these methods, we do provide an overview of the study designs frequently used in climate and health studies, available resources for more detailed information on these designs, and examples of published studies that have employed these methods. We also provide strengths and limitations to each study design.

Of note, many of these designs involve the use of aggregate or count data, and thus may be subject to the ecological fallacy (i.e., an observation made with aggregate data may not be present in individual-level data). This must be taken into consideration when making inferences based on aggregate data. Another consideration is the temporal and spatial scales of the study, which may be dependent on data availability as well as sample size issues. In some cases, data from multiple cities/states or years may need to be aggregated to have sufficient power to detect statistical associations.

4.1. Time Series Analysis

4.1.1. Background

Time series analyses identify or describe trends in longitudinal data (i.e., observations or measurements that are made sequentially in time), and may also be used to forecast or predict future occurrences of an event or outcome. Time series analysis techniques are often used in environmental and climate-related studies that consider short-term associations between health outcomes and environmental exposures. The design can be useful because of its ability to adjust for non-linear and seasonal effects, and to account for the autocorrelation (e.g., measurements/events closer in time are more correlated than measurements/events farther apart in time) inherent in time series data. However, it can be helpful to note that adjusting for confounding and autocorrelation can lead to erroneous conclusions, if not conducted or interpreted correctly.

Seasonal patterns occur cyclically (i.e., in regular intervals), such as how ED visits in Florida for asthma are the highest in winter months and lowest in the summer months. Overall trend effects consider longer-term patterns over the period of observation, such as how the overall number of ED visits for asthma has been increasing since 2005 in Florida (Figure 3). Time series techniques are able to control for these temporal patterns so that the short-term effect of the main exposure can be evaluated.[27,28]

Additionally, regular regression techniques rely on the assumption of independence between observations. Such assumptions are not appropriate for time series data, as measurements made close in time tend to be correlated. The statistical theory behind time series methods allows for valid and efficient statistical inferences.[27]

4.1.2. Study Design Specifics

The time unit of the series (e.g., hour, day, week, month) is the unit of analysis, and the outcome is a count of events (e.g., ED visits, case reports, deaths), and rarely may be continuous outcomes. Denominator data are not usually considered because population size does not typically change noticeably across the time scales used in time series analyses, particularly at small geographic areas such as a single city or county. However, if population size by sub-area within the geographic area of the analysis may affect the number of cases, then denominator data or other statistical methodology (e.g., random effects models) may be required (see *New England Heat Study* example below). Individual-level characteristics, such as age, gender, race, and smoking status, are not considered as confounders because they do not vary with the unit of analysis. Confounders that may be considered are those that vary on the same time scales as the main exposure of interest, and are known to be related to both the main exposure and the outcome of interest.[28] For example, a study assessing the relationship of daily pollen counts on ED visits for asthma may consider daily minimum and/or maximum temperature or total precipitation as potential confounders.

4.1.2.1. Analysis

The steps in a time series analysis are similar to traditional epidemiologic analyses (e.g., exploratory data analysis, modeling, model validation), with some important modifications. We briefly describe these steps here; however, each is described in more detail by Bhaskaran et al.[28]

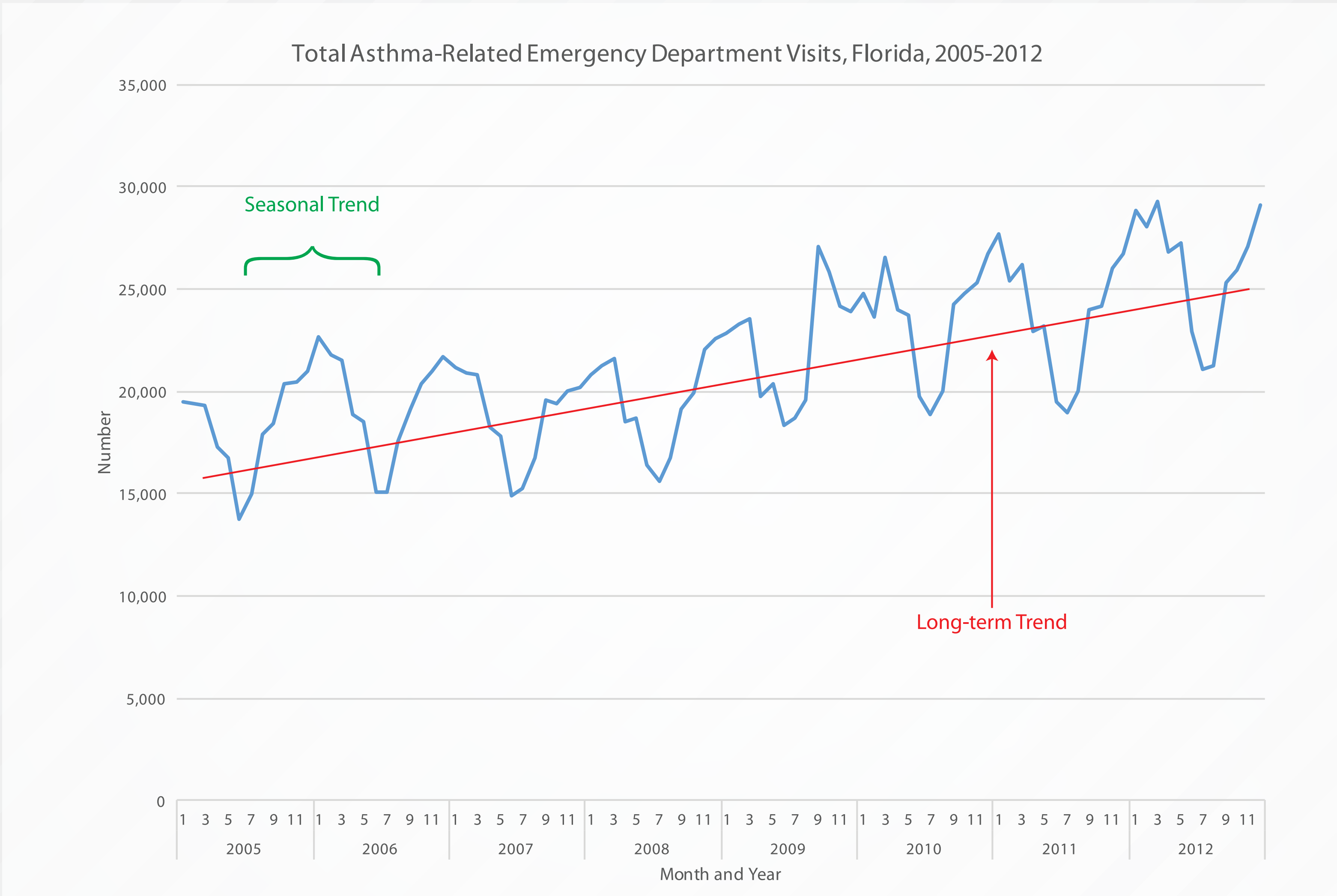


Figure 3. Time series analysis: Trend vs. seasonal effects, using monthly counts of asthma-related ED visits.

1. Exploratory data analysis

As with any epidemiologic investigation, investigators must start with exploratory data analysis to understand the data and variables. Such analyses may include summary statistics and correlation analysis for the variables of interest. With time series data, all variables of interest (i.e., exposures and outcome) can be plotted against time, using standard methods or using smoothing techniques, such as moving average plots.

2. Adjustment for seasonal and long-term trends

An important consideration in time series analysis is controlling for the confounding effects of seasonality and long-term trends. There are a variety of ways to accomplish this, including simple indicator variables for each time interval (time-stratified models), fitting sine and cosine functions of time (periodic functions), and fitting flexible spline functions (e.g., cubic splines).

3. Exposure-outcome association modeling

There are several options available for modeling exposure-outcome associations with time series data. General linear models (GLM) with Poisson distribution and log-linear models are common choices. However, for many environmental exposures, there may be delayed or non-linear effects. There may be lagged effects (e.g., the temperature today may have a greater impact on health outcomes two to three days from now) or non-linear effects (e.g., threshold effects where the rates of heat-related illness change drastically above or below a certain temperature). A way to deal with delayed exposure effects in statistical analysis is to use distributed lag linear or non-linear models. See Sections 4.1.4 and 4.1.5 for more information.

4. Model validation

Once a model has been specified to describe the exposure-outcome functions, it is common to perform a variety of model checks (e.g., plotting residuals) or sensitivity analyses to ensure that the associations remain similar under model assumptions. Such sensitivity analyses may include modifying the way that models adjust for seasonal and long-term trends, considering different lagged effects, and checking for residual confounding and other model miss-specifications.[29]

4.1.2.2. Poisson Regression Time Series

Commonly, health outcome data are aggregated temporally and spatially into counts of a given disease per day, week, or month, for a given geographic area. Poisson regression models are commonly used with count data. One example of a time series analysis uses Poisson regression models, with special extensions applied to deal with issues common in time series. Besides controlling for seasonal or long-term trends and dealing with autocorrelation, additional considerations may include addressing overdispersion, which may be considered with any Poisson regression analysis. Under the Poisson distribution, the mean of the count data is equal to its variance. Overdispersion occurs when the variance is greater than predicted (due to strong temporal patterns such as long-term trends, seasonal cycles, and outbreaks/epidemics), and adjustments to the model are needed to accurately estimate the standard errors.[28]

Several examples of using Poisson regression models are available in the published literature. One study assessed the effects of weather variability, such as monthly temperature and humidity, on the transmission of dengue fever in China using Poisson regression and generalized estimating equations.[30] Another examined the relationship between weather variability and incidence of cryptosporidiosis using Poisson regression models.[31] This study also compared Poisson regression to seasonal auto-regression integrated moving average (ARIMA) models.

4.1.2.3. Distributed Lag Models

Another example of time series analysis is a distributed lag model. This is a specialized statistical technique for understanding the relationship between variables with delayed effects (e.g., temperature). Below, a simple general linear model describes this relationship:

$$y_t = \sum_{i=1} \beta_0 + \beta_i * x_{t-i}$$

This equation demonstrates how the dependent variable at time t can be expressed as a linear function of X measured at different points (t, t-1, t-2). Time series models are commonly fit using terms for one or more lagged exposure effects. Adjusting for the different lagged effects simultaneously is known as a distributed lag model. Understanding the cumulative effects of the exposure can be estimated by summing the coefficients from a distributed lag model.[28,32]

However, the relationship between the exposure and outcome of interest may not always be considered linear. Therefore, linear and non-linear models are available. An example of a distributed lag non-linear model is described below.

4.1.3. Resources

Several available articles provide more detailed overviews and theory, as well as sample data, code, and examples. Bhaskaran et al.[28] provide an overview of the use of time series regression in environmental epidemiologic studies. The article walks through an example time series analysis using Poisson regression to assess the relationship between daily ozone levels and mortality in London from 2002 to 2006.

Sample data and related Stata and R code are provided in the online supplement to the article (http://ije.oxfordjournals.org/content/suppl/2013/05/30/dyt092.DC1).

Zeger et al.[27] describe the application of time series models to public health and biomedical data. This article provides a more in-depth discussion of the statistical theory and several more recent examples of the application of these models to a variety of public health research questions. Imai et al.[33] provide an in-depth discussion of how time series models can be applied to assess the associations between infectious diseases and weather, and Armstrong[34] discusses different modeling techniques available to describe the relationship between temperature and mortality.

New England Heat Study[32]

The distributed lag model framework explores both the non-linear effect of temperature, as well as the delayed nature of the response. Investigators used distributed lag non-linear models (DLNM package in R) to evaluate the association between heat index and daily ED admissions and deaths in seven cities in New Hampshire, seven cities in Maine, and in the state of Rhode Island between May and September from 2000 to 2010. Hourly weather station data from the NCEI were used to calculate daily maximum heat index. For New Hampshire and Maine, the geographic unit of analysis was defined as all towns within a 10-mile radius of each weather station. Daily all-cause ED visits were obtained from hospital discharge data in each state. Heat-specific ED visits were defined as cases with heat or dehydration (ICD-9 267.5; 992; E900) as a primary or secondary cause, or renal disease (ICD-9 580-589) as a primary cause. Overdispersed Poisson constrained distributed lag models controlling for long-term time trends, day of week, and federal holidays were applied to each study site. All models considered heat index over the previous 0–7 days and allowed for non-linear exposure-response functions. City-specific risk estimates were then pooled in a meta-analysis to provide a single regional estimate for all-cause and heat-related risk.

4.2. Case-Crossover

4.2.1. Background

Originally developed to study acute transient events (e.g., onset of myocardial infarction),[35] case-crossover studies are now commonly used in air pollution epidemiology[36] and are gaining popularity among climate-health researchers. The case-crossover approach is similar to a matched case-control study, with the difference being that in the case-crossover design, the cases serve as their own controls. As a result, individual-specific confounders such as underlying comorbidities or unhealthy habits like smoking, which are often not available to the analyst, can be controlled for in the analysis. It has been shown that, under certain conditions, case-crossover and time-series designs yield comparable risk estimates,[37] but the case-crossover design provides an advantage over time-series models in the examination of individual-level effect modifiers of environmental exposures.

In this study design, a person with an acute outcome is considered a case. Within a window of time around when the case is observed, the analyst chooses specific control days (or referent days) to identify records of the same health outcome for the same individual. The control days are chosen *close* to the case day such that any individual-specific factor that may be attributed to the observed health outcome would not change over that time period. The one determinant of the health outcome that could differ between the case and control days is the environmental exposure or hazard. Case-crossover model statistics summarize the strength of the relationship between the environmental exposure and the health outcome.

4.2.2. Study Design Specifics

4.2.2.1. Control (Referent) Day Selection

Choosing the window for selection of control days is a key step in a case-crossover study (Table 4). Two commonly used designs for control selection are the symmetric bidirectional and time-stratified. In the symmetric bidirectional, control days are selected as the same number of days before and after the case day. For example (see the figure below), based on the case day, control days could be 7 and 14 days before and after. In the time-stratified design, the analyst makes an *a priori* choice of a fixed time period, for example a month or half a month. Based on the case day, control days are chosen within that fixed time period for the same day of week. This way, unlike symmetric bi-directional control sampling scheme, the control days do not have a fixed temporal relationship to the case day (i.e., exchangeable). Time-stratified control selection methods are recommended to avoid "overlap bias," a type of selection bias that may occur as a result of choosing the controls based on the event time.[38]

	CONTROL	CONTROL	CASE	CONTROL	CONTROL
Symmetric Bidirectional	29-Apr	6-May	***13-May***	20-May	27-May
Time-stratified (month)		6-May	***13-May***	20-May	27-May

Table 4. Case and Control day selection in Symmetric Bidirectional and Time-stratified (month) analysis.

4.2.2.2. Data Sources

The health and environmental exposure data need to be compatible in terms of the temporal and spatial resolution. For example, in order to determine case and control days, daily observations on exposure are required. The health and exposure data also need to align spatially. For example, information used to assign a case to a location needs to be associated with the closest available source of exposure information (i.e., either some monitor-based observation or aligned with gridded modeled output).

Health data: Any acute health outcome containing data on time and location. This could be related to mortality, morbidity (e.g., hospital admission, ED visit, and physician visit), self-reported illness, etc.

Exposure data: Any transient exposure like air pollution, precipitation, pollen days, some storm events, short-term heat effects, etc.

4.2.2.3. Statistical Model

This example model uses conditional logistic regression. Functionally, the models regress the log odds of an outcome (e.g., death, hospitalization) on an exposure (e.g., daily temperature). Each matched pair is considered a separate stratum and is assigned a separate intercept which is 'conditioned' out of the analysis. The model resembles the equation below:

$$logit\ (outcome) = \alpha_0 + \alpha_{n1} STRAT_n + \beta_1(Exp) + \beta_2(Exp) + \beta_3(possible\ confounders)$$

Where:

- STRAT = stratum,
- Exp = some measurement of exposure (e.g., Max Heat Index >95°F, Ozone)
- Possible confounder = day of week

Therefore, $e^{\beta 1}$ is interpreted as the odds of death on days when the max heat index is greater than 95°F, controlling for other covariates. This estimates the risk ratio because the odds of the event (death) are rare.

Additionally, the case-crossover method allows for inclusion of interaction terms so effect modification can be evaluated.

4.2.2.4. Limitations

The time-stratified method may suffer from bias due to residual seasonal confounding.[39] Also, adjusting for overdispersion and auto correlation in counts is not possible with the time-stratified case-crossover method, though the influence of longer-term temporal patterns should be minimal because control days are typically chosen close to the case days. The conditional Poisson model has been suggested as an alternate and preferred method to account for overdispersion and auto-correlation.[40] Analysts may also consult the recent epidemiologic literature for guidance on new modifications and applications of the basic study design.

It may also be helpful to keep in mind that the case-crossover study design is most effectively used for exposures that are highly variable, such as precipitation, or exposures that have short lag times and less auto-correlation. For exposures such as heat and cold, which can have cumulative lagged effects over many days, or potentially weeks in the case of cold, the time-series methods may allow for more effective control of auto-correlation. The case-crossover design should also be used with care when examining extreme heat events. For lengthy heat waves (e.g., a 10-day heat wave), it is possible that a control day may fall within the heat wave period. In general, the analyst may want to consider which study design method can best address their specific question and whether their exposure is suited to this design.

4.2.3. Resources

There are a multiple statistical packages that can be used for case-crossover studies. For SAS, SAS Proc PHREG and Proc Logistic, as well as code examples, are available from Wang et al.[41] at http://aje.oxfordjournals.org/content/174/1/118.full.pdf+html.

For Stata and R, examples of conditional logistic model are provided in addition to those for the conditional Poisson models (the main topic) in the "Additional files" section of a study by Armstrong et al.[40] (http://www.biomedcentral.com/1471-2288/14/122).

As previously noted, there are several examples of studies using the case-crossover design to assess climate-health questions, including:

- Auger et al.,[42] Ambient heat and sudden infant death: a case-crossover study spanning 30 years in Montreal, Canada;
- Gronlund et al.,[43] Vulnerability to extreme heat by socio-demographic characteristics and area green space among the elderly in Michigan, 1990–2007 ;
- Madrigano et al.,[44] Temperature, myocardial infarction, and mortality: effect modification by individual- and area-level characteristics

4.3. Hybrid Matched Retrospective Cohort Study

4.3.1. Background

A *cohort study* is an epidemiologic study design in which subsets of a defined population are identified based on having been exposed or not exposed to a factor or factors hypothesized to influence the occurrence of a given health outcome. A *retrospective cohort study* is one in which exposures and outcomes are assessed after they have occurred.[45] A *matched cohort study* matches the exposed and unexposed groups on important confounding factors.[46] The unexposed group is selected to ensure that they are similar to the exposed group on certain characteristics, such as age, race, gender, socioeconomic factors, and even spatiotemporal variables. *Hybrid* implies that this approach is not the standard individual-level cohort study. For the purposes of this application, the unit of analysis is not an individual but an aggregate geographic unit, such as the ZIP code or county. A typical, general hypothesis for such a design is the frequency or rate of the health outcome of interest will be different in the exposed period/cohort compared to the unexposed period/cohort.

4.3.2. Study Design Specifics

4.3.2.1. Assumptions, Strengths, and Limitations

Some key *assumptions* of this particular study design are:

- Exposure occurs before disease or outcome of interest (e.g., ED visit)
- Exposed/unexposed are representative of a well-defined general population
- Unexposed group is (absence of exposure) well defined
- Outcome assessment comparable between groups
- Matching assumption: samples are not independent, must be considered in analysis

The *strengths* associated with a cohort design and applicable to the hybrid matched design include:

- A clear temporal sequence can be established (i.e., the exposure occurred before disease)
- With retrospective design, exposure and outcome have already occurred
- Multiple outcomes can be assessed simultaneously
- Matching may control for effects of measured and unmeasured confounders.

The *limitations* of this study design are:

- Subject to the ecological fallacy because aggregate data is used (i.e., hybrid)
- Typically limited to examining only one exposure
- May be difficult to define unexposed cohort.

4.3.2.2. Justification for Modeling Choice

The hybrid matched cohort study is somewhat similar to a case-crossover study in terms of the matching of different time periods for the same individual (i.e., exposed vs. unexposed period for cohort and event vs. non-event or outcome period for case-crossover), and statistical models may be similar between the two methods. This may be especially true of parameter estimates, though differences may be seen in standard errors. However, the assumptions are different across study designs in important ways.

In Florida, we chose a matched cohort study over case-crossover methods for examining the health impacts of two of our priority hazards: hurricanes and floods.[46] The objective of these analyses was to determine whether rates of ED (or hospital) visits for our health outcomes of interest were higher in the time periods immediately following a hurricane or flood event compared to a time period without such an event. The hurricane related analysis is described in the box below.

- There was no reliable individual-level data to link visits for the same patient across exposed and unexposed time periods. Therefore, a matched individual-level study was not possible, and data were aggregated to daily count by county.
- Some of the assumptions with case-crossover studies are that the exposure or event of interest is brief, and the time lag between the exposure and outcome is brief. Researchers did not feel comfortable making those assumptions with all of the health outcomes because 'exposed' periods are not transient or brief (e.g., extended clean-up periods and power outages are really the exposure for some of the health outcomes rather than the landfall of the storm itself). In general, case-crossover studies are best applied when the time lag between the exposure and disease is brief and the exposure has little carryover effects. This is not always applicable to hurricanes, but is much more applicable to short heat or precipitation events.
- Results of case-crossover studies are focused on short-term risks, and not long-term or cumulative risks. With some outcomes (e.g., mental health), those can be cumulative risks.
- Finally, study subjects (counties) were chosen based on exposure rather than outcome.

Florida Tropical Cyclone Study

The matched cohort design was used to examine the health effects of tropical cyclones (categories 3, 4, and 5 only) in Florida, occurring between 2004–2012. The unit of analysis was daily counts of ED visits in a specific county, comparing impact periods to control periods. Impact periods varied by the health outcomes considered. For example, drowning deaths related to storm surge will be more immediate (two days before to two days after) while injuries associated with clean-up can occur during evacuation and preparedness activities and days to weeks after landfall during clean-up (two days before to two weeks after). Both a pre- and post-hurricane season control period were chosen for each TC impact within the same calendar year, with appropriate 'wash-out' periods in between to limit carryover effects from previous systems. We hypothesized that rates of daily ED or hospital visits for certain health outcomes (e.g., injury, carbon monoxide poisoning) would be higher in an impact periods compared to the control periods. Associations between TCs and daily counts of ED visits were examined using conditional Poisson regression models to account for the matching (http://www.floridahealth.gov/environmental-health/climate-and-health/documents/tc-profile.pdf).

This retrospective analysis focused on the current relationship between tropical cyclones and ED visits. While there is still much to be learned about tropical cyclones and climate change, these current exposure response functions were used for disease burden projections that qualitatively assessed a change in population and change in storm intensity as the underlying assumptions.

4.3.3. Resources

A variety of statistical packages are available that can handle conditional Poisson regression models, including SAS and STATA.

In SAS, the appropriate procedure would be Proc GENMOD. The following model statement uses a Poisson distribution with log link, a population offset (log of population) to model the rate of the health outcome of interest, and a repeated subject statement where the subject is equal to the matching ID variable. A simple SAS syntax example for Proc GENMOD is provided below. Proc PHREG is also an option for conditional Poisson regression models, using Cox's partial likelihood models.

```
proc genmod data=dataset;
      class exposure_variable (ref="0") match_ID;
            model count_variable = exposure_variable/dist=poisson
link=log offset=ln_pop type3;
            repeated subject=match_ID/type=unstr;
      estimate "Exposed Period vs. Unexposed Period"
exposure_variable 1 -1/exp;
run;
```

In STATA, the CSMATCH command is available for individual-level analysis but can likely be adjusted for count-level data. Syntax example below.

csmatch depvar expvar [if expvar] [in range], group(varname) [level(.#) personvar(varlist) pairvar(varlist)]

csmatch = estimates the risk ratio for the exposure-disease relationship of interest

depvar = outcome variable, must be binary and coded as 0 (no outcome) or 1 (outcome)

expvar = exposure variable, must be binary and coded as 0 (unexposed) or 1 (exposed)

group(varname) = identifier variable (numeric or string) for the matched pairs

level(.#) = confidence level, as a fraction; default is level(.95)

personvar(varlist) = a list of potential confounding variables, must be numeric

pairvar(varlist) = a list of variables that are the same for each matched pair, must be numeric

For more information on matched cohort analysis in STATA, see: Cummings & McKnight,[46] Analysis of matched cohort data (http://ageconsearch.umn.edu/bitstream/116248/2/sjart_st0070.pdf).

4.4. Attributable Risk/Fraction

Calculating the public health burden of an exposure is a powerful epidemiologic tool. The attributable fraction (AF) is a measure that can be used to estimate the proportion of cases that can be attributed to one or more risk factors. Generically, it is also a measure of the proportion of cases that would be prevented if the risk factor(s) were completely eliminated. AF is appealing when communicating to policymakers how an intervention could lower a particular disease burden, thus lending to its increasing use among scientists interested in climate-related health outcomes.[45,47] There are two types of exposure attributable fractions (AF_e), the excess fraction and the etiologic fraction. The former answers the question "What is the proportion of outcome among those exposed that is ***attributable*** to the exposure?" The latter estimates the fraction of cases ***caused*** by the exposure and requires assumptions about the underlying biological mechanisms in addition to the epidemiologic study results.

4.4.1. Exposure Attributable Fraction (AF_e): Excess Fraction

The excess fraction captures the proportion of cases that would not have occurred in the absence of the exposure. The AF_e is most often used with climate-related exposures; it can be calculated using risk, rates, or odds making it applicable for many exposure-outcome relationships. Mathematically stated, the excess fraction formula is the risk* difference divided by the risk in the exposed (which is equivalent to the risk ratio minus one, divided by the risk ratio). The formula is defined below:

$$AF_e = \frac{R_1 - R_0}{R_1} = \frac{RD}{R_1} = \frac{R_1}{R_1} - \frac{R_0}{R_1} = 1 - \frac{1}{RR} = \frac{RR - 1}{RR}$$

where,

R0 = risk or rate in the exposed

R1 = risk or rate in the unexposed

RD = risk or rate difference

RR = risk or rate ratio.

Rates can be substituted for risks in this formula; however, the resulting AF_e will only

* Risk is also called the incidence proportion and is calculated as the number of incidence cases (in the exposed [A_1] or in the unexposed [A_0]) divided by the population ($R_1 = A_1/N$ or $R_0 = A_0/N$)

approximate the excess risk fraction if the effect of the exposure on person-time is small or if the outcome is rare over the study period. Even if the two measures are not approximate, the excess rate fraction can be used as an upper bound for the excess fraction. This property is due to the relationship between the risk ratio and the rate ratio, where the rate ratio will either approximate the risk ratio or will be further from the null. Additionally dependent on study design, the odds ratio can also be substituted for the risk or the rate ratio in the above formula, provided the odds ratio validly estimates the risk or rate ratio. For instance, in case control studies, the odds ratio will approximate the rate ratio, when the controls are sampled via incidence density sampling; the risk ratio, when the controls are sampled via case-cohort sampling; and the risk ratio, when the controls are sampled via cumulative sampling and the outcome is rare. Finally, the actual number of observed exposed and non-exposed cases can be substituted for risk in the formula. The resulting fraction is equivalent to the excess risk fraction and is often referred to as the excess case-load fraction. For the remainder of this report, excess risk and excess cases will be used interchangeably. Because the various effect measures may not always be equivalent it may be helpful to note the effect measure used in the AF_e calculation.

4.4.2. Exposure Attributable Fraction (AF_e): Etiologic Fraction

The following discussion on the etiologic fraction is abstract, in the sense that it cannot be identified with epidemiologic data. The etiologic fraction, unlike the excess fraction, is the proportion of the cases that would have occurred in the absence of an exposure, but that occurred earlier in time due to exposure.[45,48,49] When the exposure causes the outcome, those with the outcome can be divided into two groups: (1) those for whom the outcome would not have occurred if the exposure was not present, and (2) those for whom a change in the timing (i.e., earlier or later) of the outcome would have happened if the exposure was not present. Group 1 represents the "excess cases."[†] Group 2 would have the outcome regardless of exposure being present, and therefore is not an excess case (rather an "etiologic case").[50] However, as the exposure contributed to the biologic process that caused the outcome in both groups, the exposure would be etiologically relevant for both groups. Therefore, an excess case is always an etiologic case (i.e., caused by the exposure), but an etiologic case may not be an excess case. The excess risk fraction can be thought of as a lower bound for the etiologic fraction with the upper bound of the etiologic fraction being 100% (i.e., all cases are caused by the exposure).[45]

For instance, let's say we were looking at the risk of cardiovascular mortality during heat wave versus non-heat wave days over a summer. There may be a number of individuals in the population with heart conditions who, by the end of the summer, would have enough damage to their heart that a cardiovascular event would have occurred. However, the stress of maintaining thermoregulation during a heat wave among those with heart conditions may induce a cardiovascular event during the heat wave. For these individuals, the exposure, (i.e., the heat wave), caused the cardiovascular event making these individuals both an excess case and an etiologic case. On the other hand, the stress of maintaining thermoregulation during a heat wave may only increase the damage to the individuals' heart resulting in a cardiovascular event occurring outside of the heat wave, but earlier than would have occurred without

[†] Or excess rates/odds dependent on what effect measure was used in the AF equation.

the heat wave. These individuals would not be an excess case but would be considered an etiologic case (i.e., the heat wave played a causal role in the outcome).

4.4.3. Population Attributable Fraction

The population attributable fraction (AF_p) is the proportion of all the cases or rates in the population that is attributable to the exposure.[45]

$$AF_p = p_c(AF_e), where\ p_c = \frac{\#\ exposed\ cases}{\#\ total\ cases}$$

The calculation of p_c does not change regardless of the effect measure used.[49] The above formula for AF_p is valid in the presence of confounding if the effect measure (e.g., risk or rate ratio) used in the calculation was adjusted for confounding.

The attributable fraction (both AF_e and AF_p) ranges from 0 to 1, is traditionally multiplied by 100, and presented as a percent. The attributable fraction is the proportion (or percent) of the outcome cases[‡] or rate during the study period that can be attributed to the exposure. For the AF_e the proportion is among the exposed, and for the AF_p, the proportion is among the population. The attributable fraction can also be thought of in the following manner: if the exposure is completely removed from the population (or the exposed group) the number of cases or rate will eventually be reduced by the value of the attributable fraction.

> An example of the AF_e uses estimates from a study of the association between wildfires and self-reported health outcomes among Southern California students during October 2003.[51] The study reported an adjusted odds ratio of 4.42 for irritated eyes among those reporting smelling fire smoke at home indoors greater than six days compared with those reporting not smelling fire smoke. The AF_e is $^{4.42} - {}^{1}/_{4.42} \times 100 = 77.3$ which can be stated as: "it is estimated over the study period that after adjusting for baseline asthma, ethnicity, parental education, and study cohort, 77% of self-reported irritated eye cases can be attributed to smelling fire smoke." For AF_p, the World Health Organization reports that, globally, 44% of asthma cases are attributed to environmental factors that can be modified by short- or long-term interventions. Within developed countries (as defined by WHO) the prevalence of exposure in 2002 was 0.17.[52] Therefore, $0.17 * 0.44 = 7.5\%$ of the asthma burden in developed countries in 2002 was estimated to be attributed to environmental factors which can be modified by short- or long-term interventions.

Often, disease (or outcome) is the result of numerous environmental, social, biological, and behavioral risk factors that may interact with each other. The sum of the separate AF_p (or AF_e), based on each risk factor, may (and often does exceed) 100%, suggesting that there are myriad approaches for reducing the risk of disease. As the sum of different AF_p (or A AF_e) adds to infinity, the complement of the AF_p (or AF_e) cannot be calculated.[50] As such, if 15% of deaths in the study population are attributable to extreme heat exposure it would be inappropriate to conclude that (1-0.15 = 0.85) 85% is the amount that can be explained by other exposures.

> One non-climate specific example of this comes from a nested case-control study of gastroenteritis conducted in the Netherlands (1998–1999).[53] This study used

‡ Recall cases and risk are interchangeable.

incidence density sampling so that the odds ratio approximated the rate-ratio. The rate of laboratory-identified bacterial pathogens in cases was 1.35 times the rate in controls. Therefore, during the one-year period, 26% ($= {}^{1.35-1}/_{1.35} \times 100$) of the gastroenteritis rate can be attributed to bacterial pathogens. However, the rate ratio for each laboratory-identified pathogen, including *Salmonella, Campylobacter,* and Verocytotoxin-producing *Escherichia coli* were 1.43, 2.14, and 1.90, respectively, which results in an AF_e of 30%, 53%, and 47%, respectively.

The resulting value of the attributable fraction, as with other epidemiologic measures of occurrences, is highly dependent on the distribution of cofactors (i.e., effect modifiers) within the population and the incidence of competing causes (i.e., causal mechanisms that do not involve the exposure).[48,49] For instance, let's assume a hypothetically simplistic example with two similar populations, both with the same number of asthma cases and the same ozone exposure. Population A has a large proportion of smokers (i.e., a competing cause) and population B has no smokers. As a result, the number of asthma cases due to ozone exposure estimated by the attributable fraction will be higher in population B than in population A since population B does not have a competing cause. Another example: individuals without the sickle cell trait are more susceptible to a malaria parasite than those individuals with the trait. If high temperatures (e.g., 88°F) are associated with increased risk of malaria with the effect being larger in the susceptible group, then populations with a larger susceptible population will have a higher relative effect estimate (e.g., risk ratio) and higher number of cases attributable to temperature than the population with a smaller proportion of susceptible cases.

4.4.4. Attributable Number

Attributable number (AN) is a count measure that is calculated using information gleaned from the AF_e. A commonly used equation for AN comes from heat epidemiology,[54] listed below:

$$AN = \sum[AF(E_i) * MDC * ND(E_i)],$$

Where

E_i = the levels of the exposure (i.e., temperature at 70°F, 71°F, 72°F...90°F, 91°F, 92°F),

$AF(E_i)$ = attributable fraction of a particular exposure,

MDC = mean observed daily outcome (e.g., death) count, and

$ND(E_i)$ = the number of days with particular exposure

Current and future AN estimates can be used to compare the change in disease burden from one time period to the next.[54] Benmarhnia et al.[54] calculated future AN estimates of temperature-related mortality under varying climate scenarios and concluded the methodology is useful in understanding variability in climate projections and impacts on heat-related mortality. Assumptions around daily counts, changes to the populations at risk, and exposure may be identified when drawing comparisons across a time period.

4.4.5. Resources

The AF formulas discussed in this section assume dichotomous exposure. For further information about calculating AFs across multiple exposure levels or multiple exposures or calculating variance, please refer to Steenland & Armstrong.[49] For additional discussion on the common misinterpretations and misuse of AFs, please refer to Rockhill et al.[55] and Greenland.[56]

4.5. Qualitative Analyses

Qualitative studies, particularly in the climate and health field, are useful for examining the relationship between climate-sensitive hazards and associated health outcomes.[§] Qualitative studies can improve knowledge on vulnerable populations and can complement available quantitative data. Qualitative information can improve assessments of places where limited quantitative data are available, such as in rural areas. A few examples of the types of qualitative studies that are available are listed below.

4.5.1. System Assessment

The system assessment approach assesses static elements in governance systems in view of a changing climate.[57] It makes use of a combination of organized knowledge and organized power, but the interplay between knowledge and power is not often elaborated. Organized knowledge refers to the use of models or scenarios to draw plausible pictures of how the future might look. Organized power refers to the use of laws, formal organizations, climate acts, official agreements, and regulations in order to govern society with the goal of climate adaptation. Challinor et al.[58] conducted a systems assessment which uses organized knowledge (climate models, crop yield models) and organized power (regulators and regulation) to show an undesirable outlook associated with a changing climate in Africa.

4.5.2. Storytelling

Another type of qualitative study is the storytelling approach. The storytelling approach incorporates more interaction with the people who may be affected by climate-related exposures in order to determine health outcomes that may result from climate change. This may include conducting interviews with the locals, connecting with focus groups, or using more modern techniques such as interactive photo-sharing. This approach allows for the knowledge to be more personal and ethnographic. The Oregon Public Health Division used the storytelling approach to assess whether the public's concerns and solutions are heard regarding health issues affected by climate change.[59] Several examples to gather data include hosting a film screening to facilitate community conversation, surveying stakeholders and developing "story portraits," conducting interviews and documenting their stories through video, hosting a community listening session, and creating an online storybook.

§ A literature review was performed using Ebsco Host. Articles were chosen using the search terms "climate change" and "qualitative and "health". Articles were retrieved from the published dates of 1997–2015. All results (N=85) were included including academic journals, magazines, and books. In addition, expert opinion on current practices was employed to identify additional applied methods in practice currently being explored.

The focus group approach is another qualitative method that is useful in "explaining and interpreting people's lives, actions, perceptions, fears, and feelings,"[4] particularly when researching sensitive climate change issues. A focus group is a carefully selected group of people who meet to discuss a particular issue based on questions raised by a moderator. An example of a study that used the focus group approach is Tapsell et al.,[60] which conducted six focus group meetings in order to determine attitudes, stresses, behavior, and health effects caused by vulnerability to flooding in northeast England.

4.5.3. Interviews

Interviews are another approach that can be used to examine perceptions, behaviors, and opinions on certain climate change issues from community members. The interviews can either be formally structured in order to obtain specific information on certain topics, or be more open-ended in order to gather opinions and information in a more general format. One example of a study that relied primarily on interviews is Abrahamson et al.,[61] which conducted interviews to evaluate perceptions of heat wave risks to health in people over 75 living in the United Kingdom.

4.5.4. Mixed Methods

An approach that uses both qualitative and quantitative data (mixed methods) can also be used to gather knowledge. One example of such a study is Harper et al.,[62] which used qualitative and quantitative data to determine climate-sensitive health priorities in Nunatsiavut, Canada. The study included qualitative in-depth interviews with regional health representatives, qualitative PhotoVoice workshops which allow participants to take or gather photographs that reflect their ideas, thoughts, and feelings on a particular subject, and quantitative community surveys to understand community-level trends and perceptions of climate-sensitive health priorities. Participatory GIS (Geographic Information System) mapping has been used by local communities considering actions to adapt to future impacts. In areas where there is minimal data available, collecting global positioning system (GPS) and presenting the data using GIS mapping can help improve community knowledge on assessing hazards such as flood risk.[63]

4.5.5. Surveys

Surveys are helpful in providing data when the literature may not have enough evidence to link exposure and outcomes. In 2014, the Maricopa County Department of Public Health, with the assistance of the Arizona Department of Health Services and Arizona State University, evaluated the use of cooling centers during extreme heat days by providing surveys to visitors of cooling centers in Phoenix, Arizona. The visitor surveys were useful in identifying exposure risk, such as air-conditioning status at their home as well as their experience with heat-related illness. This local level information helps to provide a more accurate view of vulnerable areas.[64]

5. Connecting Exposure-Response Functions to Disease Burden Projections

Next, we present two case studies to demonstrate the interconnectedness between forecasting climate impacts and assessing vulnerabilities and projecting the disease burden. The BRACE framework was designed to be an iterative process in which end-users begin by prioritizing climate hazards and related health outcomes for targeted adaptations and public health interventions.[65] A number of hazards and climate-sensitive health outcomes are considered in BRACE Step 1, and then a ranking process based on location-specific priorities results in the disease burden being projected for a sub-set of the original health outcomes during BRACE Step 2. This process of prioritization requires a few key steps: engaging partners and stakeholders, reviewing the literature, and assessing data. There are several outputs that are assumed to be the result of these first two steps of the BRACE Framework: a vulnerability assessment that identifies hazards of interest and vulnerable populations, a measure of association for each hazard and health outcome of interest, and estimates for current and future disease burden. Similar principles would also apply for jurisdictions not implementing the BRACE Framework but working to establish location-specific exposure-response functions.

However, this process is intended to be highly flexible in order to meet the needs and reflect the resources of different jurisdictions, to be compatible with institutional priorities, and to incorporate regional environmental hazards. It can also be noted that considerations of those developing such exposure-response functions may be different based on political realities, the target audience, and intended use. However, below are two examples showing that the process of connecting BRACE Steps 1 and 2 is an iterative process that may require continuous quality improvement.

5.1. Case Study: Florida

5.1.1. Hazards Considered

Florida is vulnerable to the impacts of tropical cyclones (TCs) and experiences more landfalls than any other state. The storm surge and high winds associated with TCs contribute to a variety of direct and indirect effects on human health. The geographic impacts of storm surge and TC winds, however, may be different. Both hazards were included when the University of South Carolina Hazards and Vulnerability Research Institute (HVRI), in consultation with Florida BRACE staff, prospectively assessed the scope and severity of climate-sensitive hazards and identified the ones most likely to affect the health of people within the state (http://www.floridahealth.gov/environmental-health/climate-and-health/_documents/climate-sensitive-hazards-in-florida-final-report.pdf). Unlike sea level rise, heat, and drought, which are hazards that are easily and directly tied to General Circulation Models (GCM), the geographic scope of other hazards including storm surge and TC winds were analyzed using probabilistic models. There are synergies between these hazards, such as sea level rise intensifying the potential for storm surge. However, the Florida BRACE Program has operated with the knowledge of these possible synergies, while, choosing to describe and analyze effects from single hazards individually, in order to establish a baseline understanding of such hazards with limited resources.

5.1.2. Role of the Vulnerability Assessment

The formal hazard and vulnerability assessment, Climate-Sensitive Hazards in Florida (http://www.floridahealth.gov/environmental-health/climate-and-health/_documents/climate-sensitive-hazards-in-florida-final-report.pdf), is a large document which includes geographic analysis of hazard impacts as well as sub-county analysis of each hazard with the Social Vulnerably Index (SoVI) and a Florida-specific Medical Vulnerability Index (MedVI). The report has guided exposure-response function development and disease burden projections.

The potential impacts of tropical storm- and hurricane-force winds in Florida were calculated using Extended Best Tract data for 1988–2012 and an idealized buffer around storm tracks for 1952–1987. Return periods, or the average annual frequency of occurrence for each census tract, were calculated. The eastern coast and south Florida are at the highest risk of tropical-storm force winds. Overall, approximately 9.4 million people living in 35 counties are at high risk of tropical storm-force winds (50–75% historical, annual frequency) (Figure 4). The Panhandle and south Florida are at highest risk of hurricane-force winds. Overall, approximately 2.9 million people living in 19 counties are at high risk of hurricane-force winds (10–15% historical, annual frequency) (Figure 5).

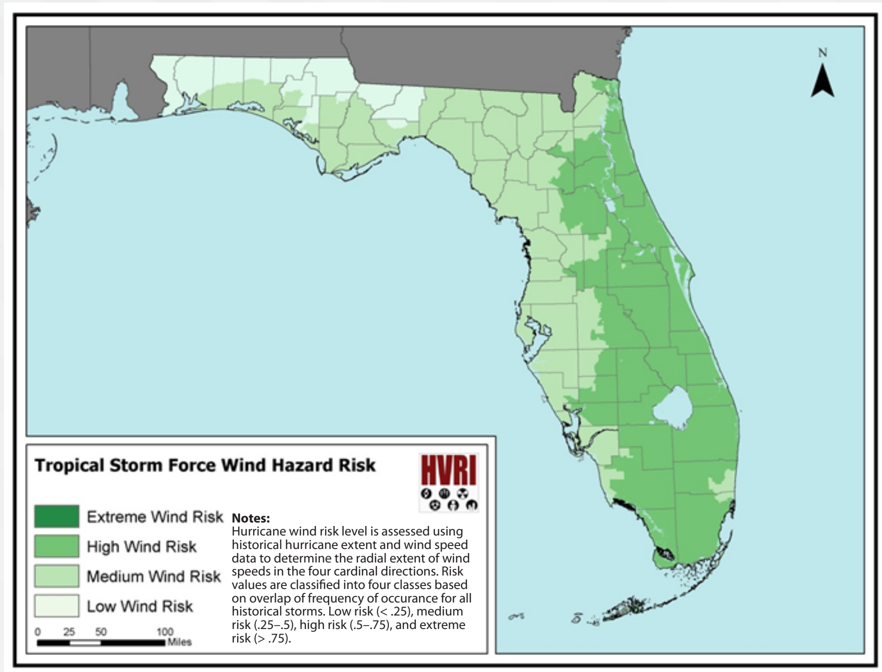


Figure 4. County-level tropical storm wind hazard risk in Florida (Source: Hazards and Vulnerability Research Institute).

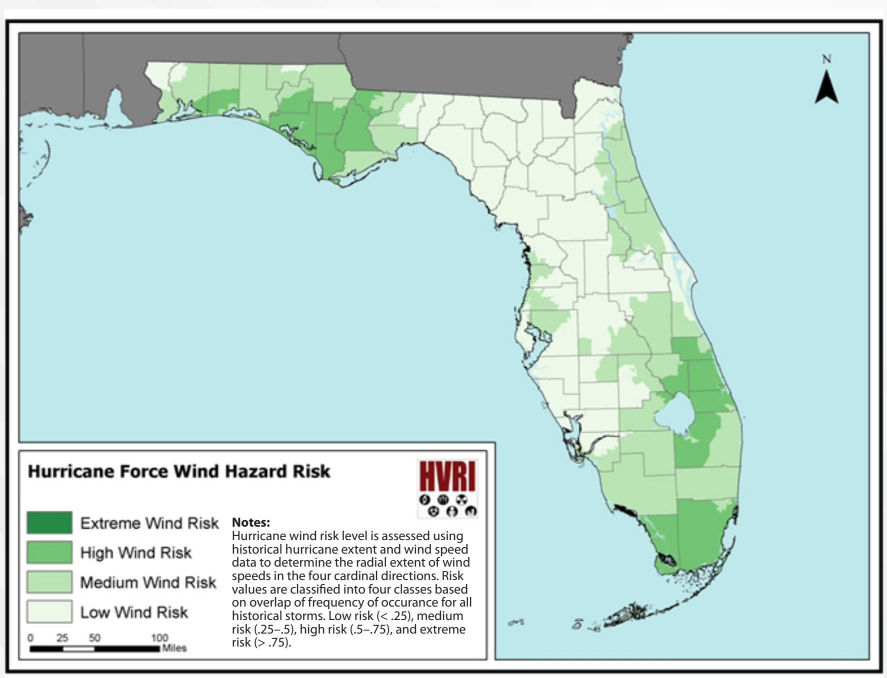


Figure 5. County-level hurricane force wind hazard risk in Florida (Source: Hazards and Vulnerability Research Institute).

The potential impact of storm surge along Florida's coastline was calculated using the NOAA Sea, Lake, and Overland Surges from Hurricanes (SLOSH) model. Depending on the direction of the storm, all of south central Florida and counties along the Gulf Coast are at highest risk of storm surge. In a Category 5 hurricane, 5.6 million people living in 38 counties are at extreme or high risk of storm surge, with half residing in Hillsborough, Lee, Miami-Dade, and Pinellas Counties (Figure 6).

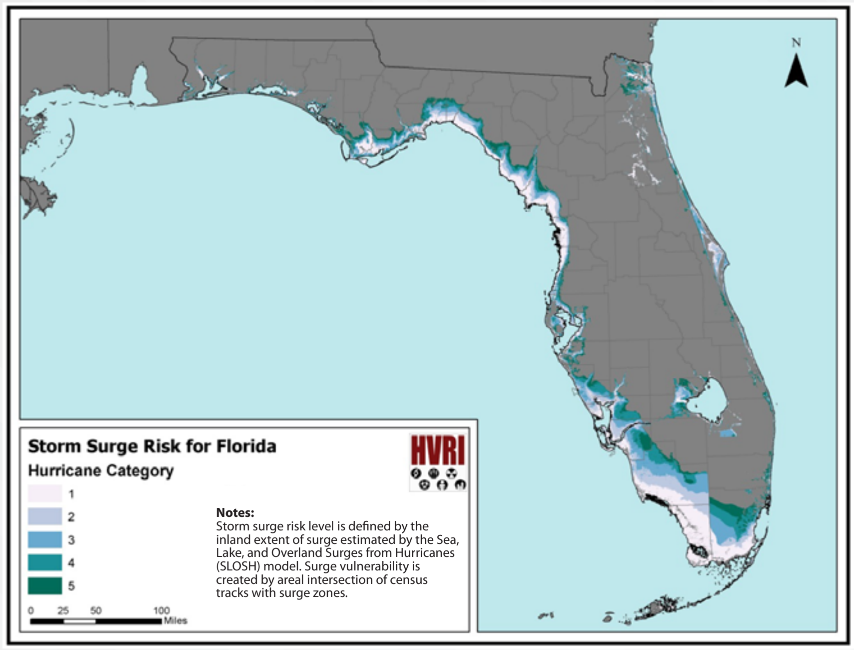


Figure 6. U.S. Census tract-level storm surge risk in Florida (Source: Hazards and Vulnerability Research Institute).

5.1.3. Stakeholder Engagement

In addition to active discussions between HVRI and FL BRACE staff during hazard prioritization and analysis, program staff worked with state and regional climatologists to confirm the appropriateness of prioritized hazards. Results of the assessment were presented to the Program's Technical Advisory Group for feedback. For TCs, additional discussions were had with faculty from the Florida State University departments of meteorology and geography. There was consensus that using GCM output for future TC projections was not appropriate as there is so much uncertainty about the validity of model output currently available.

5.1.4. Data Challenges and Solutions

Because of the uncertainty associated with TC projections in the 21st century, the Florida BRACE Program choose to utilize a qualitative approach to disease burden projections rather than quantitative for this hazard. As part of BRACE Step 1, we had developed TC-specific exposure-response functions for a variety of health outcomes, including all-cause injury and carbon monoxide (CO) poisoning, food- and waterborne diseases, and drowning. Exposure-response functions were assessed using a matched cohort study design with the unit of analysis defined as daily counts of visits or poison control center calls in a specific county. Exposure was defined as counties experiencing tropical storm-force (39–74 MPH) or hurricane-force (≥74 MPH) winds or storm surge greater than 1.2 m. Impact periods were outcome-specific. Exposure-response functions were analyzed for combined TC impacts (i.e., tropical storms and hurricanes) and separately for hurricane impacts using Poisson regression models. Rate ratios (RR) and 95% confidence intervals (CI) were calculated (Table 5). CO poisoning had the strongest associations with TC activity, with rates of CO poisoning in impact periods being 3.4 to 6.6 times the rates during control periods, and stronger associations existing when examining hurricane-only impact periods. Associations with TC impacts were also observed for injury, *Cryptosporidium, Salmonella,* and *Vibrio*. Based on the strength of associations identified in these analyses and internal discussions, we choose to project future disease burden for all-cause injury and CO poisoning. The goal of conducting these disease burden projections was to better understand and plan for the future impacts of TCs on the health of Floridians.

INDICATORS	ANY TC IMPACTS			HURRICANE IMPACTS		
	RR	LCL	UCL	RR	LCL	UCL
Injury (ED)	1.03	1.02	1.05	Did not converge		
Injury (Hospital)	1.04	1.02	1.05	1.24	1.19	1.3
CO Poisoning (ED)	3.44	2.07	5.72	11.66	5.14	26.45
CO Poisoning (Hospital)	4	2.9	5.51	10.49	5.98	18.38
CO Poisoning (FPICN)	6.59	4.48	9.7	14.94	8.2	27.22
Campylobacter	1.02	0.9	1.14	1.02	0.77	1.35
Cryptosporidium	1.26	1.04	1.52	1.26	0.61	2.62
Giardia	1.01	0.89	1.14	0.78	0.56	1.09
Salmonella	1.35	1.29	1.42	1.69	1.49	1.9
Vibrio	1.48	1.06	2.07	2.16	0.96	4.85

Table 5. Results from Florida's exposure response analysis.

Because of the interannual variability and the complexity of factors associated with TC formation, there is wide variation in **TC projections** for the 21st century. The general consensus, however, is a tendency toward increasing intensity. Therefore, projections for both combined TC impacts and hurricane-only impacts were presented, with the hurricane-only impacts representing the projected increase in intensity.

Using an attributable risk model for disease burden projections, additional pieces of information besides the exposure-response functions were needed including **baseline rates** of disease and population projections. The baseline rates of all-cause injury and CO poisoning were based on hospital data, ED data, and poison control center calls (for CO poisoning only). **Population projections** for Florida by county and year were obtained from the Florida Bureau of Economic and Business Research. Total population projections for Florida for 2020 to 2040 in ten-year increments were used.

Attributable risk models were used to project disease burden for health outcomes of interest. RRs and associated 95% CIs were converted to attributable fractions (AF) in order to obtain the estimated number of events that may occur due to TC or hurricane impacts above baseline. Some limitations to this method must be noted. The attributable risk models assume that there are no changes in baseline rates over time and no additional hurricane adaptation measures are implemented during this period. Further, direct estimates of projected changes in TC impacts were not included; instead, indirect estimates of projected increases in intensity were included by focusing on hurricane impacts only. Additional work is on-going to further improve the methods used to project TC-related disease burden among Floridians.

5.1.5. Florida Lessons Learned

Because it was the first time exposure-response functions were calculated by staff within the agency, the process was not as seamless as it could have been. For example, the TC vulnerability assessment was driven by probability-based scenarios that were not used for developing either the exposure-response functions or conducting the disease burden projections. Additionally, the complexity and uncertainty in the current climatological research surrounding TC activity in the 21st century required additional analysis planning for BRACE Step 2 and ultimately resulted in qualitative, rather than quantitative, projections.

We present a visualization of the Florida BRACE Program's process below, from the vulnerability assessment to disease burden projections, to further clarify the lessons learned in this process (Figure 7). The Program took a hazard-specific view throughout the process (i.e., examining individual rather than synergistic climate hazards and effects), beginning with a formal vulnerability assessment. A scenario-based assessment was conducted using low, mid, and high emissions GCM outputs for sea level rise, heat, and drought. A probability-based (e.g., risk indices) assessment was done for TCs, flooding, and wildland fire. Local retrospective analyses were completed for some, but not all, hazards. Such analyses were not completed for sea level rise or wildland fire based on limitations in exposure or related health data. Time series analyses were conducted for heat and drought, whereas a matched cohort-based analysis was done for both TCs and flooding. For disease burden projections, we used our Florida-specific exposure-response functions. Heat and drought-related projections were done using quantitative climate projections (i.e., based on GCM output using the A2 emissions scenario), while TCs were projected using a qualitative method. From vulnerability assessment to disease burden projections, discrepancies in methodologies arose out of necessity (e.g., qualitative projections based on intensity rather than probability indices for TCs) or based on expert advice (e.g., use of A2 admissions scenario in projections instead of the ones used in the vulnerability assessment). Because the BRACE Framework is an iterative, flexible process, the Florida BRACE program will now be able to refine activities in BRACE Step 1 based on knowledge gained conducting BRACE Step 2.

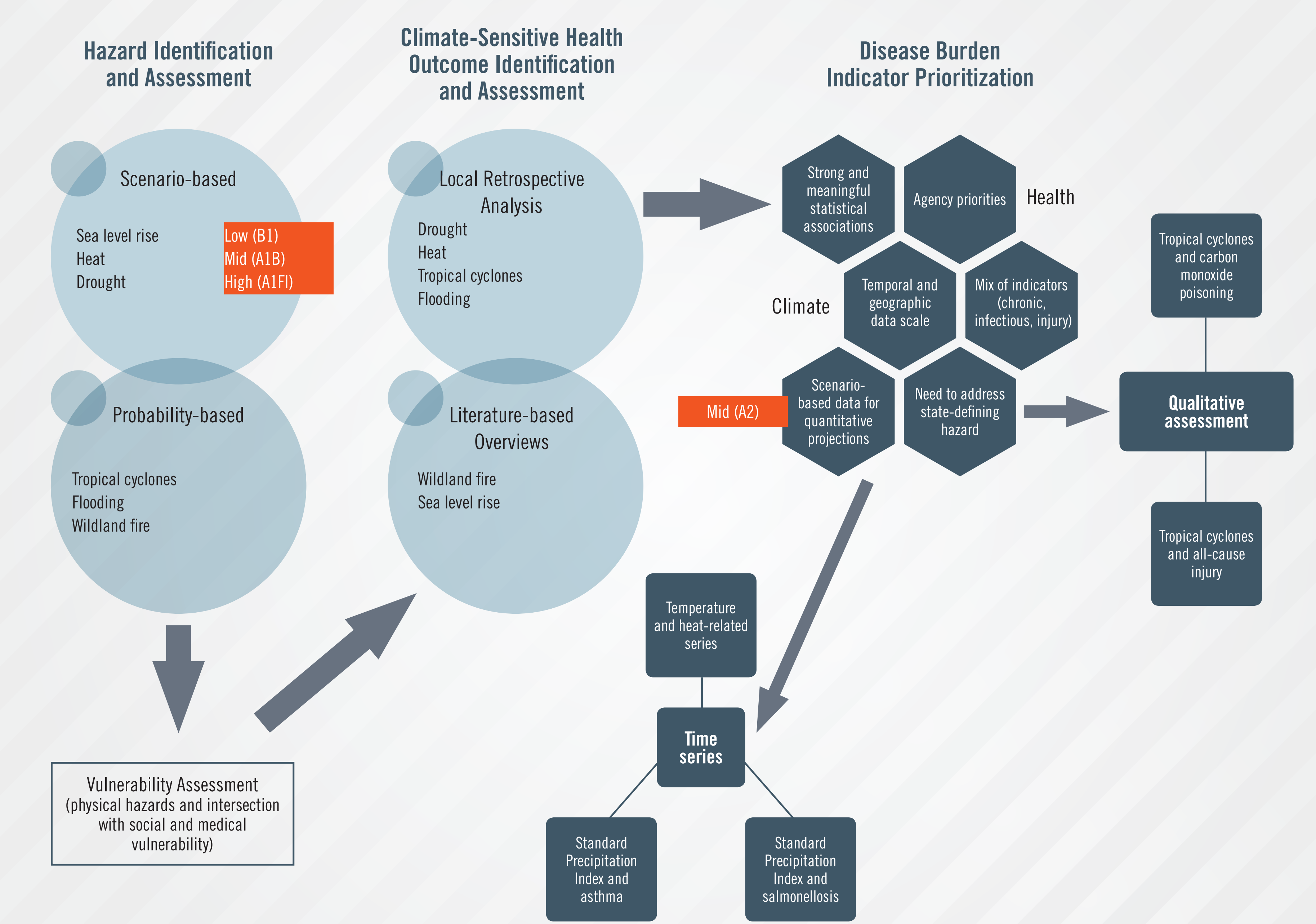


Figure 7. Visualization of Florida’s decision-making process related to BRACE Steps 1 and 2.

5.2. Case Study: New Hampshire

5.2.1. Hazards Considered

In New Hampshire (NH), the vulnerability assessment completed during BRACE Step 1 was a 3-step process. The NH Climate and Health Program evaluated social vulnerability, climate vulnerability, and hydrologic vulnerability. In these reports, relevant climate hazards were identified. A study on *Climate Change in New Hampshire* from the University of New Hampshire and *Climate Solutions New England*[66] indicated that future climate hazards include rising temperatures, a greater likelihood of days over 90°F and 95°F, and fewer days below freezing. The report also indicated more precipitation, and more extreme weather events that may result in floods. Another assessment by the United States Geological Survey (USGS) developed a watershed model that suggested NH will have greater average stream flow, increased risk of flooding, and variable groundwater recharge depending on the season.[67]

Because social factors are also important determinants of health, NH used a state-specific Social Vulnerability Index (SVI: http://nhdphs.maps.arcgis.com/home), designed by the NH Department of Health and Human Services, to determine which communities may be the most socially vulnerable to adverse health outcomes associated with climate hazards and extreme weather. The SVI is a web-based tool that allows users to visualize 15 vulnerability factors in four categories (Socioeconomic Status, Household Composition/Disability, Minority Status/Language, Housing/Transportation) at the Census tract level.

These three vulnerability assessments were then used in a prioritization process to identify the key climate hazards, associated health outcomes, and vulnerable populations to focus on during the health burden assessment in BRACE Step 2. In order to extrapolate the exposure-risk function based on future climate scenarios, the exposure metrics used in the retrospective analysis had to align with the climate metrics generated in the climate vulnerability assessment. The vulnerability assessments provide a foundation for estimating how the changes in temperature, precipitation and severe weather may affect certain health outcomes. See Figure 8 for more information on the theoretical connection between BRACE Steps 1 and 2.

5.2.2 Role of the Vulnerability Assessment

The social, climate, and hydrologic hazards identified in BRACE Step 1 provide a foundation for estimating how the changes in temperature, precipitation and severe weather may affect certain health outcomes. The major health outcomes associated with climate change were summarized in a subsequent report, *Climate Change and Human Health in New Hampshire, an Impact Assessment*.

The health burden assessment process as described by Marinucci et al.[63] links the vulnerability assessment to disease burden assessment by providing future disease burden estimates that can help public health agencies prioritize issues for future action. The primary objective of BRACE Step 2 is to determine the future disease burden associated with a changing climate. Currently, NH focuses on four health outcomes that are related to the hazards identified:

- Heat stress and heat-related illness,
- Air quality and allergy and asthma,
- Water quality and gastrointestinal (GI) illness, and
- Changing habitat and vector-borne disease.

5.2.3. BRACE Framework in Action

Local public health agencies can use the same assessment process to: (a) identify climate-related hazards and vulnerabilities and (b) link them to human health in order to quantify the future burden of disease based on climate projections. In NH, a partnership between state, regional and local health agencies is facilitating a climate and health adaptation planning process meant to build community resilience and reduce adverse health effects associated with severe weather and climate change.

Using the CDC BRACE Framework for building resilience, regional partners were encouraged to assess vulnerabilities, evaluate potential interventions, and develop adaption plans (Figure 8). To build public health workforce capacity, we created a guidebook, held trainings, and provided direct consultations to public health partners to develop new knowledge, skills and abilities related to climate adaptation (Figure 9).

So far, two local public health agencies have been funded to develop Climate and Health Adaptation Plans. Both local agencies completed written adaptation plans in partnership with existing regional Public Health Advisory Councils. A new round of funding will expand the project to additional regional public health agencies. See Figure 9 for additional information on implementation of the BRACE Framework in NH.

Figure 8. Theoretical connection between BRACE Step 1 and Step 2.

Step 1
Forecast Climate Impacts and Assess Vulnerabilities

- Contracted with academic partners at UNH to complete an assessment of vulnerabilities:
  - Created downscaled Climate Projections, and
  - Highlighted 3 key climate-related hazards: increasing temperature, increasing precipitation, and more extreme weather events.
- Contracted with USGS to complete a report on hydrologic vulnerabilities:
  - Focused on streamflow, groundwater base flow, and snowfall.
- Used existing Social Vulnerability Index to evaluate social determinates of health in the context of climate change.
- Aligned priorities with State Health Improvement Plan.
- Funded local Public Health Regions through a mini-grant process, and encouraged them to identify key climate-related hazards and vulnerable populations in specific regions through their own assessment and stakeholder engagement process.

Outcome of Step 1

Identified 3 key climate-related hazards and associated health outcomes

Step 2
Project the Disease Burden

- Conducted a literature review which is summarized in the Climate and Health Profile as well as an impact assessment written by UNH, Climate Change and Human Health in New Hampshire.
- Conducted NH-specific epidemiologic analysis focused on heat indea and heat-related illness (in partnership with the Northeast Heat Health Collaborative).
- Contracted with academic partners at Plymouth State University to create extreme weather indicators that will be used in epidemiologic analysis of extreme weather and injury.

Outcome of Step 2

Estimate the current disease rate project the burden under future climate scenarios

Figure 9. Operationalization of BRACE Step 1 and Step 2 in New Hampshire.

5.2.4. New Hampshire Lessons Learned

Successful implementation of the BRACE Framework in NH required strong partnerships with academic partners as well as local health agencies. Through a collaborative process, NH created an important set of resources that both summarize the state of the science and guide partners through the process of hazard identification, vulnerability assessment, and health burden assessment. In this work, it was helpful to choose climate hazard metrics wisely. It was important that the exposure variable used in the retrospective analysis corresponds to the exposure variable used in the climate projection. For example, NH made climate projections for daily maximum temperature, however, current epidemiologic analysis focuses on heat index. In order to project the disease burden of heat-related illness based on climate projections, assumptions will need to be made about the relationship between heat index and maximum temperature. It can also be helpful to consider the geographic resolution of analysis when conducting vulnerability assessments. It may be helpful for the chosen geography to be sensible from a decision-making perspective, and align with municipal boundaries. Overall, implementing the BRACE Framework has been an iterative process involving significant stakeholder input at both the state and local level.

6. Conclusions

This document serves as a resource for public health practitioners interested in applying the CDC five-step BRACE framework to plan for and adapt to climate hazards. We provide examples of relevant environmental and health data that can be used in climate and health analyses; resources of study designs and statistical methods commonly used to assess relationships between climate hazards and health effects; and methods for quantifying future climate-related disease burden. Common methodologies include case-crossover studies, time series analyses, and matched cohort studies.

Ultimately, the value of developing strong exposure-response functions and connecting retrospective analysis to disease burden projections is to create a strong foundation for planned adaptation activities. Laying the groundwork for adaptation planning can be helpful for public health decision makers. However, this foundation may look different in every jurisdiction. Possible outcomes include a single number quantifying risk for a single exposure and outcome, a comparative assessment, or a qualitative assessment. Regardless of format, it may be helpful if public health practitioners collaborate with the appropriate partners and develop analyses that are locally meaningful and can be used to inform decision making.

7. Acknowledgments

Thank you to everyone whose dedication has made this effort successful, including members of the BRACE Methods CoP, state and local public health agencies that support climate adaptation and collaboration within the field, and the CDC Climate and Health Program for providing funding and technical assistance to support implementation of the BRACE Framework.

Special thanks to Kathleen (Katie) Bush (NH), Sharon Lee (MA), Neil Muscatiello (NY), Chaity Naik (CDC), Nathaniel (Nate) Schafrick (VT), and Jared Ulmer (VT) for being CoP reviewers for this document. Your fresh eyes and excellent comments significantly improved the final product.

For more information on the Climate-Ready States and Cities Initiative grantees, please see: http://www.cdc.gov/climateandhealth/crsci_grantees.htm. The following software was used for analysis and displaying results: Geographic Information System (GIS) ArcMAP v.10.0 (ESRI: Redlands, CA); SAS v9.3 (SAS Institute: Cary, NC); and R (R Development Core Team: Vienna, Austria).

Recommended Citation: Kintziger KW, Jagger MA, Conlon KC, Bush KF, Haggerty B, Harduar-Morano L, Lane K, Roach M, Thie L, Uejio CK, BRACE Methods Community of Practice. 2017. *Technical Documentation on Exposure-Response Functions for Climate-Sensitive Health Outcomes*. Climate and Health Technical Report Series. Climate and Health Program, Centers for Disease Control and Prevention.

8. Works Cited

1. Kingston, F., Lennox & Addington Public Health Informatics, *Workshop on Syndromic Surveillance of Health and Climate-Related Impacts: Workshop Summary Report*. 2014.
2. Yao, J., M. Brauer, and S.B. Henderson, *Evaluation of a wildfire smoke forecasting system as a tool for public health protection.* Environ Health Perspect, 2013. 121(10): p. 1142-7.
3. Davis, R.E., G.R. McGregor, and K.B. Enfield, *Humidity: A review and primer on atmospheric moisture and human health.* Environ Res, 2016. 144(Pt A): p. 106-16.
4. American Meteorological Society. *Glossary of Meteorology*. 2013; Available from: http://glossary.ametsoc.org/wiki/Main_Page
5. Karl TR, Williams CN, and Q. FT, *United States Historical Climatology Network (HCN) Serial Temperature and Precipitation Data*. 1990, Oak Ridge National Laboratory, Climate and Environmental Sciences Divsion, Carbon Dioxide Information Analysis Center.
6. Hughes P, et al., *United States Historical Climatology Network Daily Temperature and Precipitation Data*. 1992, Oak Ridge National Laboratory, Climate and Environmental Sciences Divsion, Carbon Dioxide Information Analysis Center.
7. Easterling DR, et al., *United States Historical Climatology Network (US HCN) Monthly Temperature and Precipitation Data*. 1996, Oak Ridge National Laboratory, Climate and Environmental Sciences Divsion, Carbon Dioxide Information Analysis Center.
8. Easterling DR, et al., *United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data for 1871-1997*. 1999, Oak Ridge National Laboratory, Climate and Environmental Sciences Divsion, Carbon Dioxide Information Analysis Center.
9. Network, U.S.H.C. 2013, Oak Ridge National Laboratory, Climate and Environmental Sciences Divsion, Carbon Dioxide Information Analysis Center: Oak Ridge, Tennessee.
10. Lawrimore, J.H., et al., *An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3*. Journal of Geophysical Research: Atmospheres, 2011. 116(D19): p. n/a-n/a.
11. Menne, M.J., et al., *An Overview of the Global Historical Climatology Network-Daily Database*. Journal of Atmospheric and Oceanic Technology, 2012. 29(7): p. 897-910.
12. NOAA National Centers for Environmental Information, *Global Historical Climatology Network (GHCN)*. 2016: Asheville, North Carolina.
13. NOAA National Climatic Data Center, *Quality Controlled Local Climatological Data*. 2010.
14. NOAA National Centers for Environmental Information, *Quality Controlled Local Climatological Data (QCLCD)*. 2016: Asheville, North Carolina.
15. N, L. *The quality control of the integrated surface hourly database*. in *Proceedings of the 84th American Meteorological Society Annual Meeting*. 2004. Seattle, Washington USA.
16. Information, N.N.C.f.E., *Integrated Surface Database (ISD)*. 2016: Asheville, North Carolina.
17. National Weather Service, *Weather Service Operations Manual B-17 (revised),* C.P. Management, Editor. 1987, NOAA-NWS: Silver Springs, Maryland.
18. Daly, C., et al., *Observer Bias in Daily Precipitation Measurements at United States Cooperative Network Stations*. Bulletin of the American Meteorological Society, 2007. 88(6): p. 899-912.
19. Information, N.N.C.f.E., *Cooperative Observer Network (COOP)*. 2016: Silver Springs, Maryland.
20. Information, N.N.C.f.E., *Global Surface Summary of the Day (GSOD)*. 2016: Asheville, North Carolina.
21. Goodall, J.L., et al., *A first approach to web services for the National Water Information System*. Environmental Modelling & Software, 2008. 23(4): p. 404-411.
22. Survey, U.S.G., *Water Data for the Nation*. 2016.
23. National Atmospheric Deposition Program, *Maps and Data*. 2014: Champaign, Illinois.
24. Mixon Hill, *Clarus system user guide,* U.S. Department of Transportation Federal Highway Administration, Editor. 2011: Overland Park, Kansas.

25. Needham, H.F. and B.D. Keim, *A storm surge database for the US Gulf Coast.* International Journal of Climatology, 2012. 32(14): p. 2108-2123.

26. Rosenzweig, C. and W. Solecki, *New York City Panel on Climate Change 2015 Report Introduction.* Annals of the New York Academy of Sciences, 2015. 1336(1): p. 3-5.

27. Zeger, S.L., R. Irizarry, and R.D. Peng, *On time series analysis of public health and biomedical data.* Annu Rev Public Health, 2006. 27: p. 57-79.

28. Bhaskaran, K., et al., *Time series regression studies in environmental epidemiology.* International Journal of Epidemiology, 2013. 42(4): p. 1187-1195.

29. Flanders, W.D., et al., *A method for detection of residual confounding in time-series and other observational studies.* Epidemiology, 2011. 22(1): p. 59-67.

30. Lu, L., et al., *Time series analysis of dengue fever and weather in Guangzhou, China.* BMC Public Health, 2009. 9(1): p. 395.

31. Hu, W., et al., *Weather variability and the incidence of cryptosporidiosis: comparison of time series poisson regression and SARIMA models.* Ann Epidemiol, 2007. 17(9): p. 679-88.

32. GA, W., et al., *Heat-Related Morbidity and Mortality in New England: Evidence for Local Policy.* Environ Health Perspect, 2017. In Press.

33. Imai, C., et al., *Time series regression model for infectious disease and weather.* Environ Res, 2015. 142: p. 319-27.

34. Armstrong, B., *Models for the relationship between ambient temperature and daily mortality.* Epidemiology, 2006. 17(6): p. 624-31.

35. Maclure, M., *The case-crossover design: a method for studying transient effects on the risk of acute events.* Am J Epidemiol, 1991. 133(2): p. 144-53.

36. Carracedo-Martinez, E., et al., *Case-crossover analysis of air pollution health effects: a systematic review of methodology and application.* Environ Health Perspect, 2010. 118(8): p. 1173-82.

37. Lu, Y. and S.L. Zeger, *On the equivalence of case-crossover and time series methods in environmental epidemiology.* Biostatistics, 2007. 8(2): p. 337-44.

38. Janes, H., L. Sheppard, and T. Lumley, *Overlap bias in the case-crossover design, with application to air pollution exposures.* Stat Med, 2005. 24(2): p. 285-300.

39. Whitaker, H.J., M.N. Hocine, and C.P. Farrington, *On case-crossover methods for environmental time series data.* Environmetrics, 2007. 18(2): p. 157-171.

40. Armstrong, B.G., A. Gasparrini, and A. Tobias, *Conditional Poisson models: a flexible alternative to conditional logistic case cross-over analysis.* BMC Med Res Methodol, 2014. 14: p. 122.

41. Wang, S.V., et al., *Potential for bias in case-crossover studies with shared exposures analyzed using SAS.* Am J Epidemiol, 2011. 174(1): p. 118-24.

42. Auger, N., et al., *Ambient Heat and Sudden Infant Death: A Case-Crossover Study Spanning 30 Years in Montreal, Canada.* Environ Health Perspect, 2015. 123(7): p. 712-6.

43. Gronlund, C.J., et al., *Vulnerability to extreme heat by socio-demographic characteristics and area green space among the elderly in Michigan, 1990–2007.* Environ Res, 2015. 136: p. 449-61.

44. Madrigano, J., et al., *Temperature, myocardial infarction, and mortality: effect modification by individual- and area-level characteristics.* Epidemiology, 2013. 24(3): p. 439-46.

45. Rothman KJ, Greenland S, and L. TS, *Modern Epidemiology.* 3rd ed. 2008, Philadelphia, Pennsylvania: Wolters Kluwer Health/Lippincott Williams & Wilkins.

46. Cummings P and M. B, *Analysis of matched cohort data.* Stata J, 2004. 4(3): p. 274-281.

47. Gasparrini, A. and M. Leone, *Attributable risk from distributed lag models.* BMC Med Res Methodol, 2014. 14: p. 55.

48. Greenland, S. and J.M. Robins, *Conceptual problems in the definition and interpretation of attributable fractions.* Am J Epidemiol, 1988. 128(6): p. 1185-97.

49. Steenland, K. and B. Armstrong, *An overview of methods for calculating the burden of disease due to specific risk factors.* Epidemiology, 2006. 17(5): p. 512-9.

50. Poole, C., *A history of the population attributable fraction and related measures.* Ann Epidemiol, 2015. 25(3): p. 147-54.

51. Kunzli, N., et al., *Health effects of the 2003 Southern California wildfires on children.* Am J Respir Crit Care Med, 2006. 174(11): p. 1221-8.

52. Pruss-Ustun A and Corvalan C, *Preventing disease through healthy environments: Towards an estimate of the environmental burden of disease,* World Health Organization, Editor. 2006: Geneva, Switzerland.

53. de Wit, M.A., et al., *Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology.* Am J Epidemiol, 2001. 154(7): p. 666-74.

54. Benmarhnia, T., et al., *Variability in temperature-related mortality projections under climate change.* Environ Health Perspect, 2014. 122(12): p. 1293-8.

55. Rockhill, B., B. Newman, and C. Weinberg, *Use and misuse of population attributable fractions.* Am J Public Health, 1998. 88(1): p. 15-9.

56. Greenland, S., *Concepts and pitfalls in measuring and interpreting attributable fractions, prevented fractions, and causation probabilities.* Ann Epidemiol, 2015. 25(3): p. 155-61.

57. Vink, M.J., A. Dewulf, and C. Termeer, *The role of knowledge and power in climate change adaptation governance: a systematic literature review.* Ecology and Society, 2013. 18(4).

58. Challinor, A., et al., *Assessing the vulnerability of food crop systems in Africa to climate change.* Climatic Change, 2007. 83(3): p. 381-399.

59. Oregon Climate and Health Program, *Story projects and listening projects: Methods for engaging diverse communities.* 2016.

60. Tapsell, S.M., et al., *Vulnerability to flooding: health and social dimensions.* Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 2002. 360(1796): p. 1511.

61. Abrahamson, V., et al., *Perceptions of heatwave risks to health: interview-based study of older people in London and Norwich, UK.* J Public Health (Oxf), 2009. 31(1): p. 119-26.

62. Harper, S.L., et al., *Climate-sensitive health priorities in Nunatsiavut, Canada.* BMC Public Health, 2015. 15(1): p. 605.

63. Singh, B.K., *Flood Hazard Mapping with Participatory GIS: The Case of Gorakhpur.* Environment and Urbanization Asia, 2014. 5(1): p. 161-173.

64. Hondula DM, et al., *The role of cooling centers in climate adaptation: Observations from the 'Phoenix Heat Relief Network Evaluation Project', in Proceedings of the Third Symposium on Building a Weather-Ready Nation: Enhancing Our Nation's Readiness, Responsiveness, and Resilience to High Impact Weather Events.* 2015: Phoenix, Arizona.

65. Marinucci, D.G., et al., *Building Resilience against Climate Effects—A Novel Framework to Facilitate Climate Readiness in Public Health Agencies.* International Journal of Environmental Research and Public Health, 2014. 11(6).

66. Wake C, et al., *Climate change in southern New Hampshire: Past, present and future.* 2014, Climate Solutions New England, Sustainability Institute, Univeristy of New Hampshire: Durham, New Hampshire.

67. Bjerklie DM, Ayotte JD, and C. MJ, *Simulating hydrologic response to climate change scenarios in four selected watersheds of New Hampshire.* 2015, United States Geological Survey,. p. 53.