

Health Impacts of Extreme Heat and Wildfire Smoke Evaluation Framework¹

Created By: Karie Nicholas, Foundation for Health Care Quality and work group sub-committee members - Jess Heit, Kelly Naismith, Karie Nicholas, Kristina Petsas, Christopher Chen, Asa Sakrison, Colin Field, Sarah Halvik, Mike Myint, Onora Lien, Petra Eichelsdoerfer, Krista Lynch, Donya Karimi, Luann Chen, Amanda Avalos, Susanne Quistgaard, Nicole St. Clair, Stephan Wheat.

¹ Adapted from the Department of Industry, Science, Energy and Resources Evaluation Framework and [Evaluation Framework | Better Evaluation](#)

[February 14, 2025]

Health Impacts of Extreme Heat and Wildfire Smoke Guidelines

This evaluation framework provides an overall framework for evaluations across different organizations within the Washington State health care system that contribute to the mitigation of and response to extreme heat and wildfire smoke events.

This evaluation framework includes:

- **definitions and key concepts**
- **principles and standards**
- **Information on resources to help align evaluations across system actors**
- **guidelines for setting priorities on what, when and ways to evaluate**
- **Health System roles and responsibilities.**

Document administration

Version history

Version	Date	Description	Author
1.0	February 15, 2025	Original draft by subcommittee	See title page

Contents

Health Impacts of Extreme Heat and Wildfire Smoke Evaluation Framework	1
Document administration	3
1. Background and Overview	6
1.1 Introduction	6
1.2 Guideline overview	6
2. Metrics Alignment	7
3. Types of Evaluations	10
3.1 Process and program evaluations	11
3.2 Evaluation of monitoring and surveillance activities	15
3.3 Impact evaluations	16
4. Evaluation Alignment	17
4.1 Guideline logic	17
4.2 Evaluation questions	18
4.3 Evaluation Matrix	18
4.4 Data Matrix Template	19
5. Methodology	19
5.1 Roles and responsibilities	19
5.2 Ethical Standards and Cultural Considerations	21
5.3 Common Contextual Factors	22
5.4 Timelines	22
5.5 General Methods	24
5.6 Risks and limitations	24
Appendix A Theory of Change	27
Appendix B Data Collection Matrix	28

Glossary

Accountable Communities of Health - a neutral convener, coordinating body, investor, and connection point between the health care delivery system and local communities. (Washington State Health Care Authority, 2024)

Audience – In Bree reports, an audience is a category of “system-actors”. For example, a common audience is “health plans” and a common system-actor would be a specific insurance company.

Care-variation - differences in process of care across multiple clinics, areas, patient groups, insurance types, etc. (Bree Collaborative).

Concordance of care – Organizational and individual activities, interactions, policies and procedures that have a high degree of alignment with best practice recommendations (i.e. for the purposes of this framework best practices are considered to be the Bree Collaborative Guidelines). (Bree Collaborative)

Equity/Equity Lens - A just outcome that allows everyone to thrive and share in a prosperous, inclusive society. (Propel Alanta, 2024) A way of viewing, analysing, or evaluating data that takes vulnerable, disadvantaged, or small groups of people into consideration to assure that all outcomes and impacts are equal (Bree Collaborative).

Evaluation - determination of the value, nature, character, or quality of something. (Merriam-Webster, 2024) A systematic determination and assessment of a subject's merit, worth and significance, using criteria governed by a set of standards. (Wikipedia, 2024)

Guideline – an action to improve health care for a specific health care service

Health Ecosystem - a complex network of all the participants within the healthcare sector. It is a community that consists of patients, doctors, and all the satellite figures who play a role in the medical care received by the patient or their hospital stay. This can include service providers, customers, and suppliers. Recently, the healthcare ecosystem has grown to include electronic health entities and virtual care providers. (Definitive Healthcare, LLC, 2024)

Implementation - the translation of guidelines into practice.

Public Employees Benefits Board (PEBB) Contracts - medical and dental plans that provide health benefits to 222,000 public employees and retirees. (Washington State Health Care Authority, 2024)

Report – A report is multipage document on a health care service

School Employees Benefits Board (SEBB) Contracts - medical, dental, and vision plans that provide health benefits to more than 130,000 employees of the state's school districts and charter schools, as well as union-represented employees of the nine educational service districts. (Washington State Health Care Authority, 2024)

System-actor – A specific type of organization that participates in health care in some way. Example: X health insurance company, the Washington State Department of Health, a specific provider, etc.

1. Background and Overview

1.1 Introduction

This Evaluation Framework outlines future evaluation activities that can be used to measure the outcomes and impacts of the Bree Collaborative *Health Impacts of Extreme Heat and Wildfire Smoke* Guidelines during the life cycle of the report. This report for evaluation planning has been developed by the Bree Collaborative Sub-committee of the *Health Impacts of Extreme Heat and Wildfire Smoke* Workgroup and the Bree Collaborative's Measurement and Evaluation Program Manager.

This document details the evaluation framework within which the future evaluation[s] of the implementation of these guidelines may be conducted. Establishing this framework early in your organizations guideline implementation life cycle ensures that the programs developed from Bree guidelines are prepared for future evaluations and helps instill an evaluative mindset from the outset. The framework provided by this document should be referred to during the implementation process and used to inform the drafting of an evaluation plan by each organization. It is recommended that it be reviewed periodically or in response to significant program, regulatory, or environmental events.

While this framework is expected to inform the evaluations outlined herein, the evaluations themselves may deviate from this framework based on input from various stakeholders and the program's evaluative needs at the time of each evaluation. This document is meant to provide alignment across multiple audiences for the purpose of comparison and to facilitate state-wide measurement on the progress and outcomes of the adoption of the Bree guidelines.

The framework provides guidance for different types of evaluations at different levels across the healthcare ecosystem. It details the reasons behind recommendations for particular types and timings of evaluation activities, makes recommendations for types of evaluations by audience, identifies domains for the development of evaluation questions, and identifies the data which should be available, or which will have to be collected to answer these questions.

Because an iterative process of system improvement and monitoring and surveillance are parts of the recommendations of the *Impacts of Extreme Heat and Wildfire Smoke* guidelines, all organizations that adopt the guidelines should consider some form of evaluative system or a process or program evaluation, at a minimum.

This framework has been prepared by taking into account the strategic importance of the guidelines and the expected level of resourcing for evaluations at each organization.

1.2 Guideline Overview

A **Bree Report** is defined as *a multipage document on a health care service, identified by Bree members as needing improvement that provides information and guidelines for actions different audiences can take within the health care ecosystem to improve the health of that chosen report topic*. A report may also be referred to as an **intervention** for the purposes of evaluation. A **Bree Collaborative Guideline** (previously called a recommendation in earlier

Bree reports) is defined as *an action to improve health care for a specific health care service*. Reports include multiple guidelines for many different system-actors.

The *Health Impacts of Extreme Heat and Wildfire Smoke* report was developed by the Bree Collaborative in 2024 and published in January of 2025 (also called Bree guidelines or guidelines). The purpose of this report is to provide guidance on health ecosystem coordination and response improvements in order to reduce the impacts to patients during heat or wildfire smoke events. The aim of mitigating the effects of heat and smoke downstream is to reduce impacts to human health and to save health care dollars and resources during and after a heat or smoke event.

These guidelines were submitted to the Washington State Health Care Authority for the purpose of implementation as part of their Medicaid and other contracting activities with the intention of improving financing, prescribing and refill delays, outreach before an event, and changing EMS utilization rates, expanding access to mitigation activities and resources, identifying those at risk during an event, and providing a guidance on evaluating performance after an event for continuous improvement efforts. The report was also published to the Bree Collaborative website for the purpose of implementation by Bree Collaborative members and by health care providers, purchasers, payors and community partners in general, in Washington State and beyond.

The components of this intervention (categories of activities that most or all of the system actors are expected to take) are **preparation and response, measurement and communication alignment, education/workforce training**, increases in **risk assessment standardization and risk mitigation, service coordination and service interruption mitigation**, changes to **financing**, and **infrastructure and capacity development**.

The Bree guidelines apply to multiple system actors (clinicians, health systems, health plans, public health, etc.) that play a part in mitigating or responding to heat or wildfire smoke events, from the family level through the state level. The goals of the Bree guidelines are to 1) minimize morbidity and mortality of extreme heat and wildfire smoke in Washington state, 2) prepare patients and health systems to respond to extreme heat/wildfire smoke through infrastructure building, education and awareness, community engagement, and establishing robust monitoring and early warning systems, 3) respond swiftly to extreme events through targeted outreach, allocating resources and clinical protocols, and 4) improve prevention and response through long-term surveillance and iterative improvement processes.

2. Metrics Alignment

The Bree Collaborative's *Health Impacts of Extreme Heat and Wildfire Smoke* guidelines report includes surveillance, response monitoring and iterative improvement protocols as a foundational component of the intervention that supports all other recommendations.

It is essential that organizations adopting these guidelines measure response and outcomes in the same or similar ways. This section provides information on which metrics are recommended and general information on how to align surveillance and monitoring efforts. How to use these measures and metrics for different types of evaluations is defined further down in this report.

Audience specific objectives and component goals can be found in the Evaluation Matrix found in section 2.7 which will help organizations build out process and program evaluation and surveillance and monitoring programs.

Surveillance and monitoring alignment

The report recommends that all organizational audience types (i.e. all except patients and families) monitor risk and response as part of the intervention. Heat related risks and air quality risks are defined in Appendix A of the report. As part of the risk mitigation component of this intervention, organizations should use these definitions to develop monitoring systems and to inform evaluations after an event.

Populations at risk during events are defined in Appendix A of the report and these definitions should also be used as part of the risk mitigation component of an evaluation or of a surveillance program.

Surveillance and monitoring evaluations should use the same before/during/after structure as recommended in the guidelines to monitor and assess the effectiveness of their actions in each phase of an event response.

Heat/smoke risks –

- National weather service forecast for HEAT a 24-hour period: 0=little to no risk from expected heat; 1= minor; 2= Moderate; 3= Major; 4=Extreme
- Environmental Protection Agency AQI: Good=0-50; moderate=51-100; unhealthy for sensitive groups 101-150; unhealthy=151-200; very unhealthy=201-300; Hazardous=301 and higher.

Health risks

- **Disabilities** – any person with mobility, communication, or self-care limitations.
- **Chronic conditions** – (N18.XX Chronic Kidney Disease; I20-25 Ischemic Heart Disease; I30-52 Other forms of heart disease; E08-E13 Diabetes Mellitus; J45 Asthma, J46 status asthmaticus; J44 COPD; J40-47 chronic lower respiratory disease; J60-70 lung diseases due to external agents; J30-39 other diseases of the upper respiratory tract; J80-84 other respiratory disease principally affecting the interstitium)
- **Pregnancy** – (O09 supervision of high-risk pregnancy; 010-016 edema, proteinuria, and hypertensive disorders in pregnancy, childbirth, and the puerperium; O80—O82 encounter for delivery)
- **Behavioural health** – SUD, psychotic disorders, other mental health disorders, developmental disorders, intellectual disorders
- **Medications** – (see Appendix D of the *Health Impacts of Extreme Heat and Wildfire Smoke* Guidelines report for a complete list)

Social risks

- **Demographics** - Infants and children, pregnant women, older adults (age 0-18 and ages 65+)
- **Activities** - Athletes, outdoor and some indoor workers, emergency responders
- **Social need** - People experiencing homelessness, people with low income (problems related to house and economic circumstances Z59); people who are incarcerated; people who rent, marginalized communities (problems related to social environment)

Z60); Other - education and literacy Z55; problems related to employment and unemployment Z56; Occupation exposure risk Z57; Problems related to upbringing Z62; problems related to primary support group Z63; problems related to certain psychosocial circumstances Z64; problems related to other psychosocial circumstances Z65)

State-wide structural measures alignment

These measures may be calculated differently depending on the organization. The intent of these measures is to identify gaps in preparedness. Organizations are encouraged to use the Bree Collaborative's [Survey Question Bank](#) and Survey Bank to share measurement methodologies for these measure concepts.

- Increase cooling centre capacity (individual center capacity, number of cooling centers per capita, distance to cooling centers, number of counties or LHJs with adequate per capita cooling center capacity)
- Number of organizations using early warning system to inform their programs, stratified by audience type (see section 3.1 for definitions)
- Increase in prevalence of or access to preventive items (air conditioners, air filters, cooling centres, etc.), stratified by geography and/or population at risk
- Increase number of organizations with an up-to-date, guideline aligned emergency preparedness plan

Capacity Measures

These measures should be calculated similarly for each organization. AHRQ has provided aggregation of over 700 relevant [ICD-10 codes](#) to ~530 categories which will help align capacity measurements. The intent of these measures is to capture additional stresses on emergency services and emergency departments during a heat event. This data can be used to quantify the impacts and costs of heat events.

- ED visits/ICU visits that include (CPT and ICD codes for altered mental status, respiratory failure, hyperthermia, intubations and ventilator use, or non-invasive ventilation, invasive and non-invasive cooling measures,) AND during an identified heat event or smoke event (see heat/smoke risk above) *(See evaluation type recommendations for more information on inclusion/exclusion criteria and timeframes for measurement.)*
- ED diversions OR EMS call incidents during identified heat or smoke event (see heat/smoke risk above)

Morbidity and mortality measures should use a timeframe of start of event to \leq 10 days after event and/or weekly trends compared to historical trends. Other measurement methods can be submitted to the Bree Collaborative for inclusion in the Implementation Guide or Evaluation tools.

Morbidity measures alignment

- *ED visits for respiratory illness.*
 - *Recommended conditions and codes include: COPD, Myocardial Infarction, Stroke, and Acute Kidney Injury AND involving heat exposure based on key words in triage notes OR chief complaints fields OR (exposure to fire/smoke X010XX, X011XX, X018XX;dehydration E860, P741; J70 respiratory*

conditions due to smoke inhalation; Natural (force of nature) X398XX; Smoke and poor air quality Z77.110, X08.8; Disaster Z655; Occupational exposure Z5739, Z576) AND an identified heat event (as defined by the report).

- *For more information on measurement methods, codes, queries, definitions, etc. Please visit the Bree Collaborative **Measurement Bank**.*
- *Occupational health incidents: for more information please visit the Bree Collaborative **Measurement Bank**.*
- *ED visits or hospitalization for heat stress conditions*
 - *Recommended conditions include: Heat stroke, heat exhaustion, heat cramps, heat syncope, heat rash, Rhabdomyolysis, stratified by zip code, race and ethnicity, age, occupation, homelessness, disability status, incarceration status and chronic conditions (See surveillance and monitoring section above) AND involving environmental heat exposure based on key words in triage notes or chief complaints fields AND an identified heat event.*
 - *For more information please visit the Bree Collaborative **Measurement Bank**.*

Mortality measures alignment

- *Heat Stress Deaths: resident deaths having any underlying or contributing ICD-10 code as follows: X30; T67.0-T67.9; Cases with a code of W92 (man-made source of heat) anywhere in the record are excluded: deaths are included only if they have occurred between May and September for each year.*
 - *Stratification: census tract or census block, race, ethnicity, age, pregnancy status, homelessness, disability status, incarceration status, and occupation (Note: low power is a concern for this measure and organizations should enumerate their methods for data masking or deidentification of numbers)*
 - *More information on how to align measures can be found in the Bree Collaborative **Measurement Bank**.*
- *Excess deaths during heat events: difference between expected deaths and actual deaths; from time of event start to ≤ 10 days after event.*
 - *Stratification: census tract or census block, race, ethnicity, age, pregnancy status, homelessness, disability status, incarceration status, and occupation (Note: low power is a concern for this measure and organizations should enumerate their methods for data masking or deidentification of numbers)*
 - *More information on how to align measures can be found in the Bree Collaborative **Measurement Bank**.*

3. Types of Evaluations

This framework provides guidance for the types of evaluations (e.g. process, monitoring and impact) that will assist in the demonstration of the usefulness of the Bree Guidelines. Organizations may also use this framework to improve their process of care, identify pinch-points or lessons learned, assess outcomes of changes made, monitor state-wide progress on the goals of the guidelines, and/or determine the impact of guidelines adoption on their patients' health, workforce, costs, etc.

Bree Collaborative's *Health Impacts of Extreme Heat and Wildfire Smoke* includes surveillance, response monitoring, and iterative improvement protocols as a foundational component that supports all other recommendations. As such, each organization should be implementing some type of ongoing evaluative system, the results of which can be used for

the purposes of demonstrating the outcomes of the Bree guidelines adoption or leveraged in a broader evaluation of the impact of the guidelines adoption across the state.

As equity is an important part of the Bree Collaboratives' work and a primary focus of the *Health Impacts of Extreme Heat and Wildfire Smoke* guidelines, strategies for assessing equity in access, care, and outcomes should be included in all types of evaluations.

For organizations that do not have evaluation programs or that do not routinely conduct evaluations, more information on evaluation planning and execution can be found at: [Evaluation.gov](https://www.evaluation.gov) | [Evaluation 101](#)

3.1 Process and Program Evaluations

It is proposed that type of evaluation be conducted by all relevant stakeholders except families and patients

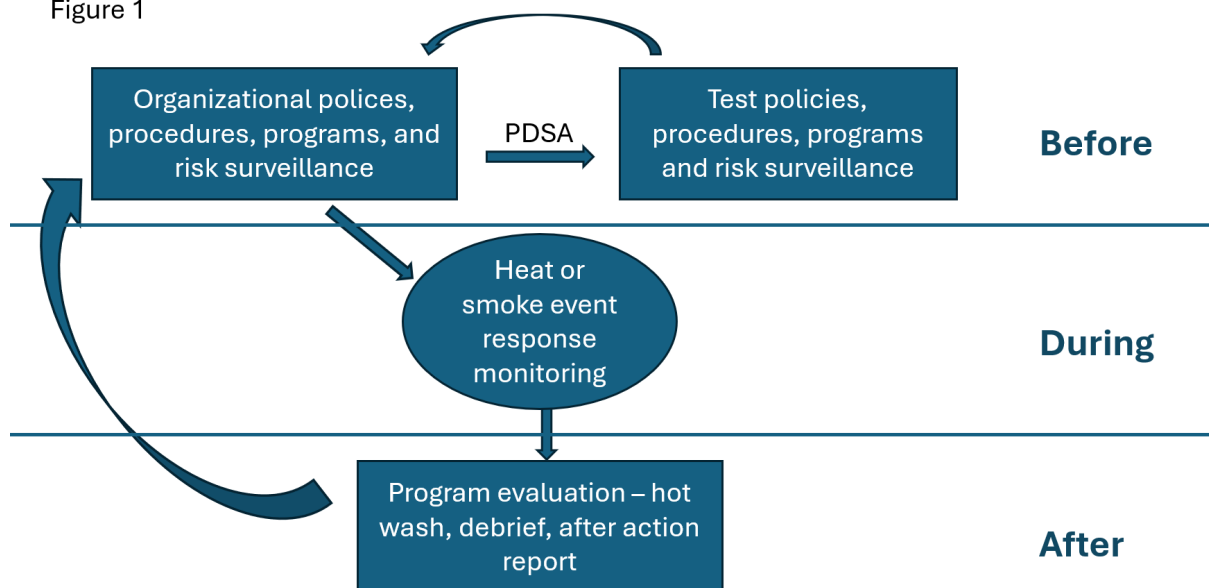
Process evaluations focus on implementation details, describing a program's services, activities, policies, and procedures. These types of evaluations can answer questions such as "Is the program reaching its intended participants?" or "How are inputs contributing to program functioning?"

Program evaluations assess final outcomes, determining whether a program achieved its goals. This type of evaluation can answer questions such as "Did participants experience the desired outcomes?" or "What changes were made to improve the quality of the program?"

The patient outcomes and impact the Bree guidelines hope to achieve are to reduce morbidity and mortality during heat or wildfire smoke events, as defined by the outcome and impact metrics, through enhanced and coordinated preparation and response. Process and program evaluations help illustrate how or if changes to processes recommended by the guidelines improve the outcomes and impacts.

In this intervention, process and program evaluation work are integrated and are an essential component and should occur before, during, and after an event. See figure 1 below.

Figure 1



Generally speaking, process evaluation focuses on implementation details, describing a program's services, activities, policies, and procedures and program evaluations assess final outcomes, determining whether a program achieved its goals. Process evaluations usually focus on the initial implementation (e.g. when teams are being created, infrastructure is being built, etc.) of a program to allow decision makers to identify early issues regarding program administration and delivery and take corrective action if necessary. Program evaluations usually focus on outcomes to determine if an intervention program is effective.

Organizations should leverage PDSA cycles as part of their processes of response and capacity improvement work. Objectives by audience (long term care, health plans, etc.) and by component can be found in the Evaluation Matrix in section 2.7 and PDSAs can be useful tools to improve processes designed to meet certain objectives.



PDSA resources:

[Institute for Healthcare Improvement \(IHI\)](#)

[Agency for Healthcare Research and Quality \(AHRQ\)](#)

[The Deming Institute](#)

Strong recommendations for developing an evaluation process:

- Use [Bree score cards](#) to align policies, procedures and activities with recommendations
- Use [Bree Survey Question Bank](#) to align survey and evaluation questions across multiple stakeholders
- Use Bree Collaborative [Measurement Bank](#) to align process and program measures
- Use structural measures, as appropriate
- Use Evaluation Matrix to identify objectives for each component (see section 2.7)
- Include an equity perspective process evaluation planning (see section 2.1)
- Use NWS [Heat Risk Tool](#) and [AirNow](#) to develop processes for monitoring event occurrences and warnings

This framework assumes that organizations involved in direct patient care (health systems, hospitals, and direct care teams) will have an established quality improvement program and that they will include a heat and smoke process, structure, and outcomes component in their existing quality improvement work, which should serve the same purpose as process and program evaluations and inform the organizations about the effectiveness of their program.

Organizations not directly involved in patient care should consider a program evaluation plan to be a key part of their implementation work to determine if their activities meet the objectives and goals for each component of this intervention (section 1.2) as outlined in the Evaluation Matrix (section 2.7).

Organizations should include the following steps in their program evaluation and/or process evaluation (see section 2.2) as recommended by FEMA:

- **Hot Wash** - A Hot Wash provides an opportunity for event response team and participants to discuss response strengths and areas for improvement immediately following the event. The Hot Wash should be led by a facilitator who can ensure that the discussion remains brief and constructive. The information gathered during a Hot Wash can be used during the After-Action Report/Improvement Plan (AAR/IP) process, and future event suggestions can be used to improve future responses or preparation. Hot Washes also provide opportunities to distribute community/response team feedback forms, which, when completed, can be used to help generate the AAR/IP.
- **Debrief** - Immediately following the event, a short debriefing should be conducted with event response team members to ascertain their level of satisfaction with the response, discuss any issues or concerns, and propose improvements. Planners should collect response attendance lists, provide copies to the response planning leaders, collect community/response team feedback forms, and develop debriefing notes.
- **"Controller/Evaluator" Debriefing** - The C/E Debriefing provides a forum for response planners and evaluators to review the event. The response planning team can facilitate this debriefing, which provides each planner and evaluator with an

opportunity to provide an overview of the functional area they observed and to discuss both strengths and areas for improvement. During the debriefing, planners and evaluators can complete and submit their response team feedback forms. Debriefing results are captured and may be included in the AAR/IP. This debriefing provides a forum for facilitators and evaluators to discuss strengths, areas for improvement, and progress in completing event response objectives.

- **After-Action Report (AAR)** - The After-Action Report is the document that summarizes key information related to evaluation. The length, format, and development timeframe of the AAR depend on the event type and scope. The main focus of the AAR is the analysis of core capabilities. Generally, AARs also include basic event information, such as the risk level, type of event (smoke, heat or both) dates, location, affected organizations, geographic area(s) in which the event occurred, specific threats or hazards (i.e. loss of key infrastructure such as pharmacies, roads, etc.), a brief scenario description, and the names and roles of the response team members. The AAR should include an overview of performance related to each event response objective and associated core capabilities, while highlighting strengths and areas for improvement. Therefore, evaluators should review their evaluation notes and documentation to identify the strengths and areas for improvement relevant to the participating organizations' ability to meet event response objectives and demonstrate core capabilities.
- **Improvement Plan** - Upon completion, the evaluation team provides the draft AAR to the organizational leadership or administration, who determine which areas for improvement require further action and distribute improvement plans to appropriate teams or sections. Areas for improvement that require action are those that will continue to seriously impede capability performance if left unresolved. As part of the improvement planning process, leadership should identify corrective actions to bring areas for improvement to resolution and determine the team or section with responsibility for those actions.

Table 1: resources for iterative evaluation

Hot wash	Debrief/Evaluator Debrief	After-action report/improvement plan
FEMA Hot-wash template Public Health Ontario Hot-wash Guide Why and How to Hot-wash After a Crisis	Types of critical incident debriefings	After-action report/improvement plan template After-action report/improvement plan template (homeland security)

Strong recommendations for evaluating program effectiveness:

- Assess organizational ability to develop a timeline for program evaluations that aligns with heat or smoke events (i.e. annually, bi-annually or greater if using a hot-wash strategy)
- Assess outcomes and impacts from the inclusion of measurements of patient and family/community education
- Include response time measurements in your assessments
- Evaluate your program with an equity lens

- Assess the use of the risk level definition in Appendix A of the guideline report or use the [Washington Tracking Network](#) to define severity of events during evaluation

Soft recommendations:

- Conduct a cost/benefit analysis for your program
- Organizations with an educational focus in their intervention (i.e. public awareness, employee/patient education, etc.) include measurements of the success of their initiatives or trainings

3.2 Evaluation of Monitoring and Surveillance Activities

It is proposed that this evaluation be conducted by: public health agencies, regional response networks, health plans, health systems, outpatient clinics, urgent care, other applicable audiences (See Table 2 below).

This evaluation type should focus on evaluating change to monitoring areas of risk and variation in response during an event. The aim of evaluating changes to your monitoring events is to determine the impact that guideline recommendations have on the ability of audience to conduct these activities.

The Bree Collaborative *Health Impacts of Extreme Heat and Wildfire Smoke* report specifically recommend that health plans, health systems, and outpatient clinics monitor *hospital and ED visits*, and that emergency departments and urgent care participate in the *documentation of employment status* for the purpose of monitoring heat and smoke exposures at the state level.

Patient risk for vulnerability to smoke should be (included, conducted) by health plans, health systems, primary care and outpatient clinics such as paediatrics or perinatal care providers.

The report further recommends that all audience types implement a protocol or system to monitor for *smoke and heat events* for the purposes of preparation.

Evaluations should be focused on answering questions about how the implementation of Bree guidelines affected changes to organizations ability to conduct surveillance and monitoring activities, increased or changed alignment of these activities with other organizations efforts, and impacted the ability of organizations to better identify risks and/or individuals at risk.

Table 2: Monitoring methods by audience

Type of monitoring	Methods	Audience type
Event surveillance	Dashboards, protocols	All
Patient risk surveillance	Patient registries, dashboards	Primary care, outpatient care, health systems, health plans

Utilization surveillance (Northwest Health Care Response Network - WATrac)	Dashboards, evaluations	Health systems, health plans, urgent care and emergency departments, EMS
Outcome and Impact surveillance	Dashboards, evaluations	Health systems, outpatient clinics, public health

Strong recommendations:

- Assess the use of NWS [Heat Risk Tool](#) and [AirNow](#) to implement measures for event occurrences and warnings into environmental surveillance protocols or dashboards
- Assess the use of 1) standardized patient work information and/or 2) patient registry and 3) patient health information to identify patients at risk and to create triggers for
- Assess the outcomes and impacts of WA DOH and WA HCA sharing claims data for monitoring and surveillance purposes

Soft recommendations:

- Organizations conducting monitoring and surveillance should assess the impact of participating in state-wide work to share data.

3.3 Impact Evaluations

It is proposed that this evaluation be conducted by state agencies or affiliates, health plans, health systems.

An impact evaluation relies on rigorous methods to determine the changes in outcomes which can be attributed to a specific intervention based on cause-and-effect analysis. (American University, 2024)

The Bree Collaborative aims to improve the quality of patient care, patient outcomes and affordability in Washington State, to that end, the measurement of the impact of guidelines adoption may be undertaken by select system actors across Washington State.

The Washington State Health Care Authority and DOH, in partnership with the Bree Collaborative, are planning an impact evaluation, however other organizations such as other state agencies, Accountable Communities of Health, large health systems or health plans, may also consider conducting an impact evaluation as it pertains to their own organization or service area.

Other organizations should participate, as requested, in impact evaluations conducted by the Bree, HCA and DOH and share any process or program evaluations they have conducted to help clarify level of adoption of the Bree guidelines.

Organizations listed above that wish to conduct their own impact evaluations should assess guideline adoption impact on one or more of the following measures:

- *Heat Stress Deaths, stratified by race, ethnicity, age, pregnancy status, homelessness, disability status, incarceration status, and occupation (Note: low power is a concern for this measure and organizations should enumerate their methods for data masking or deidentification of numbers) OR Excess deaths during heat events*
- *ED visits for respiratory illness, COPD, Myocardial Infarction, Stroke, and Acute Kidney Injury (See section 2.1)*
- *Occupational health incidents (See health risks in section 2.1)*
- *Heat Stress Hospitalization, stratified by zip code, race and ethnicity, age, occupation, homelessness, disability status, incarceration status and chronic conditions (See section 2.1)*

Impact evaluations should seek to compare impact to a prediction of what would have happened (a counterfactual) in absence of guideline adoption.

Because the purpose of the Bree is to increase quality, address variations in care and reduce health care costs, organizations that conduct impact evaluations may want to include a cost benefit analysis in their evaluation plans.

Strong recommendations:

- Include an equity lens in impact evaluations
- Include a care-variation lens in impact evaluations
- Use Bree score cards to measure concordance of care within each organization
- Use the Bree Collaborative [Measurement Bank](#) to align impact measures.
- Measure impacts at the census level using WA DOH categories (see Measurement Bank)

Soft recommendations:

- Include cost benefit analysis in impact evaluation planning

4. Evaluation Alignment

4.1 Guideline Logic

At the heart of each guideline is a ‘theory of change’ (Appendix A) by which workgroup members determine the outcomes sought and how that change can be achieved across the healthcare ecosystem. This theory of change describes how the implementation of the Bree Guidelines contributes to a chain of results flowing from the buy-in, resource utilization and capacity building, to affect medium to long-term outcomes that result in an impact for all patients in Washington State.

The concepts underpinning this guideline report are that increasing system actor’s *a priori* knowledge about heat and smoke events and education on appropriate preparation and response, changes to data sharing and reporting, and expansion of resources for response will increase system-actors’ abilities to identify individuals at risk and respond in a timely and comprehensive manner during an event, resulting in a reduction of adverse events due to heat or wildfire smoke.

System-wide logic Model can be found in Appendix A.

The Bree Collaborative offers evaluation resources, including our [Evaluation Tool Depot](#), to assist with the development of logic models at the organizational level. Organizations logic models can focus evaluation questions on outcomes and processes of interest that are appropriate for their services. They can clarify the policy and program intentions and clarify alignment between activities and objectives.

4.2 Evaluation Questions

Across the lifetime of these guidelines, evaluations need to include a range of questions that promote accountability, address gaps in care, and promote learning from system-actors experiences.

The Bree has identified four main domains for systems transformation in our [Roadmap to Health Ecosystem Improvement](#) which can be used to help develop evaluation questions which are appropriate to inform the effectiveness and impact of our guidelines: **equitable care, integrated/holistic care, data usability and transparency, and financing**. In addition to these “pillars of transformation”, the roadmap identifies levers of change which can also be used to develop evaluation questions. They include **clinical workflows, transparent reporting, education, patient engagement, coordination, contracts and networks, legislation and regulation, organizational policy changes, and data infrastructure**.

To support alignment, the Bree Collaborative has developed a [Survey Question Bank](#) which can be used to share evaluation questions across organizations participating in evaluation. Although still in its infancy, the Question Bank can be built out by participants through submission of their research questions or survey questions. Organizations may also draw from the question bank to help develop evaluations that are comparable across multiple organizations, sectors, areas, or populations.

Evaluation questions for each evaluation type can be developed to align with this roadmap and with the guideline logic and should form the basis of an evaluation plan and the Terms of Reference.

Note that not every evaluation should address all the evaluation question domains or all of the levers of change – they should be spread out across different stakeholder organizations, or across different types of evaluations such as monitoring and impact evaluations.

4.3 Evaluation Matrix

The Bree has created an evaluation matrix to align audience specific recommendations with audience specific objectives, component specific goals, and recommended metrics to measure success for each component, including recommended data sources so that guideline components can be measured in a common manner.

The Evaluation Matrix can be found [HERE](#).

4.4 Data Matrix Template

The Bree has created a data matrix template to help organization plan their data collection for evaluation purposes.

The Data Matrix Template can be found in Appendix B.

5. Methodology

This section discusses roles and responsibilities across the health care ecosystem, ethical standards, common contextual factors, timelines for collaboration, methodologies, risk for evaluation success and limitations of evaluation activities

5.1 Roles and Responsibilities

The Bree defines implementation as the “translation of guidelines into practice”. For the purposes of evaluation, we are interested in how organizations translate our guidelines into their own context or setting and what the results of their implementation are.

The Bree uses the term “Audiences” or “System-actors” in place of the term “stakeholders” for clarity. There may be one or many different organizations within an audience category (for example, there will be multiple “health plans” but only one Washington State Department of Health) or there may be multiple audiences within a single organization (for example, a health system, it’s associated clinics or hospitals and the clinicians). Finally, some organizations may play more than one role (for example, the HCA is both a purchaser and a government agency, or a health system may choose to evaluate both its patient care activities and the purchasing for its employees’ health insurance plans).

There are many system-actors or audience types with roles in implementing and evaluating the Opioid Use Disorder Treatment Guidelines across Washington State in order to affect and measure changes to care processes, financing, and outcomes across the health care eco-system. These are:

- Washington State Agencies
 - Health Care Authority (HCA)
 - » Accountable Communities of Health
 - Department of Health (DOH)
 - Local public health jurisdictions
 - Department of Social and Health Services (DSHS)
 - Legislature
- Health plans
- Health care purchasers/employers/Unions trusts
- Health Care
 - Clinicians
 - Outpatient Clinics/Urgent Care Clinics

- Hospitals
- Community organizations
 - Patients and families
 - Long-term care facilities
 - Home Health Agencies
- First responders/EMS

Table 4.1.1 below outlines broad roles and responsibilities for system-actors with regard to the *Health Impacts of Extreme Heat and Wildfire Smoke* guidelines. Bree collaborative score cards allow organizations to measure their current level of adoption of recommendations for specific activities or actions. For example, any employer that has implemented the Bree guidelines should evaluate the extent to which their organizations have implemented the recommended supports for individuals in the work environment (universally promote employee understanding of behavioural health benefits, universal communication around services offered, behavioural health-related components in employee wellness programs, Reduce employment barriers).

Organizations may have more than one role in the implementation and evaluation of the guidelines. For example, all systems actors will want to evaluate their primary function role (e.g. hospitals providing patient care) as well as their secondary roles as employers. The roles and responsibilities summarized in table 4.1.1 can help organizations define the success of their implementation across the multiple roles they play within the health care ecosystem.

Table 4.1.1: Roles and responsibilities in the health care ecosystem

Each organization has different roles and responsibilities as system-actors within a health care eco-system that provides quality care to patients. The roles and responsibilities of different organizations as defined by these guidelines are outline in the table below:

System actor role	Responsibility
State organizations	MCO purchasing (see health plans section) Risk and outcomes surveillance, monitoring, and dissemination Public education and guidance Legislative funding changes Program evaluations
Local Public Health Jurisdictions	Infrastructure prep and response planning Community coordination Public education and communication before and during events
Health Plans	Providing adequate coverage for patients at risk Enhanced support during events Outreach and education Program evaluations

Purchasers/employers/Unions	Develop requirements for plans that are purchased (e.g. benefits design) Support and protect workers during an event Employee training Event risk monitoring
Health Systems and providers	Infrastructure prep and response planning Adverse events monitoring and response System-wide coordination/data sharing Event risk monitoring Patient risk identification, documentation, and mitigation Patient education
Hospitals	Screening and triage Infrastructure prep and response planning Patient care Data Sharing
Community Organizations	Infrastructure prep and response planning Communication and education before and during events Event risk monitoring Patient risk monitoring
Emergency Services	Train staff Event risk monitoring Document and report heat or wildfire smoke related illness

Each evaluation should be overseen by a governance body established by the organization. It is not within the scope of this framework to define how each individual organizations evaluations should be governed; however, this framework sets out some general information, in sections 3.2 through 3.5, for governance bodies to consider and for organizations to consider when establishing their governance body. At a minimum, the governance body should include representation by the program's policy and delivery teams. Observers or subject matter experts from other areas should also be invited to participate as required.

As part of their evaluation plan, organizations should consider including a table, similar to table 4.1.1 above, of internal roles and responsibilities as part of their evaluations which include who is responsible for the following: *Agree to the Terms of Reference and evaluation plan, provide feedback on the evaluation report, chair of the governance group to sign off on the final evaluation report, provide evaluation guidance and input to evaluation plan, draft the evaluation Terms of Reference and evaluation plan for the evaluation; Conduct, manage, or advise on evaluation activity as required; Provide program data and guidance on program administration and delivery as required; and Provide data and input as required.*

5.2 Ethical Standards and Cultural Considerations

Equitable care is one of the pillars of the Bree Collaborative's *Roadmap to Health Ecosystem Improvement* and, as a matter of course, the Bree Collaborative encourages all implementation and subsequent evaluation work to consider an equity lens. Organizations

may refer to the Foundation for Health Care Quality's [web page](#) for further guidance when planning an evaluation.

Use of an IRB, patient safety considerations, HIPAA requirements, and other common research ethical standards should be reviewed and applied where necessary.

As noted in section 2, small numbers can be problematic in the impact measures. Organizations should consider statistical methods appropriate for the analysis of small numbers and/or data masking.

Strong recommendation:

Organizations should include equity considerations for one or more of the following groups in their evaluation plan: Rurality, race and ethnicity, disability status, homeless status, incarceration status, and individuals with chronic conditions.

5.3 Common Contextual Factors

Organizations should consider, at a minimum, the following contextual factors when planning their evaluations:

- **Washington State geography** – rugged terrain, land use, event risk
- **Demographics** – cultural, age distribution, language, health factors (i.e. chronic conditions)
- **Governmental relationships**– tribal, federal, international
- **Existing infrastructure** – **internet accessibility** defined by Washington State Office of Broadband, roads and bridges, ferries, public buildings, etc.
- **Job distribution** – outdoor work, indoor work, work from home

5.4 Timelines

The *Health Impacts of Extreme Heat and Wildfire* recommend process evaluation as one element of the infrastructure and capacity development component of this intervention. The appropriate timeline for these process evaluations is after an event defined by the guidelines has occurred. The Bree Collaborative recommends using section 2.2 above to develop process evaluations that are aligned with the guidelines and appropriate for system actors and that organizations designing these process evaluations consult with the Bree, if necessary, during the planning phase. It is anticipated that the majority of process evaluation work will take place during the fall and winter or directly after heat or smoke events and organizations should plan evaluations accordingly.

The Bree is not recommending that these process evaluations be submitted to the Bree, however, organizations that are leaders in implementation or are implementing in unique ways may want to consider a case study with the Bree Collaborative.

Figure 4.2.1 outlines the general sequence of events for program, monitoring, impact evaluations. There are three points at which organizations should consider collaboration with the Bree Collaborative: during the evaluation planning process, during the initial data collection process, and to submit a copy of the final evaluation.

Organizations may also consider closer partnerships with the Bree for evaluation support, or with the Foundation for Health Care Quality, for leveraging data from other programs within

the Foundation such as OB COAP, CBDR, or Smooth Transitions. In such cases, organizations may want to adjust their evaluation timelines to align with the Bree's awards or reporting initiatives or with FHCQ programs data collection schedules.

Figure 4.2.1: Roles and responsibilities

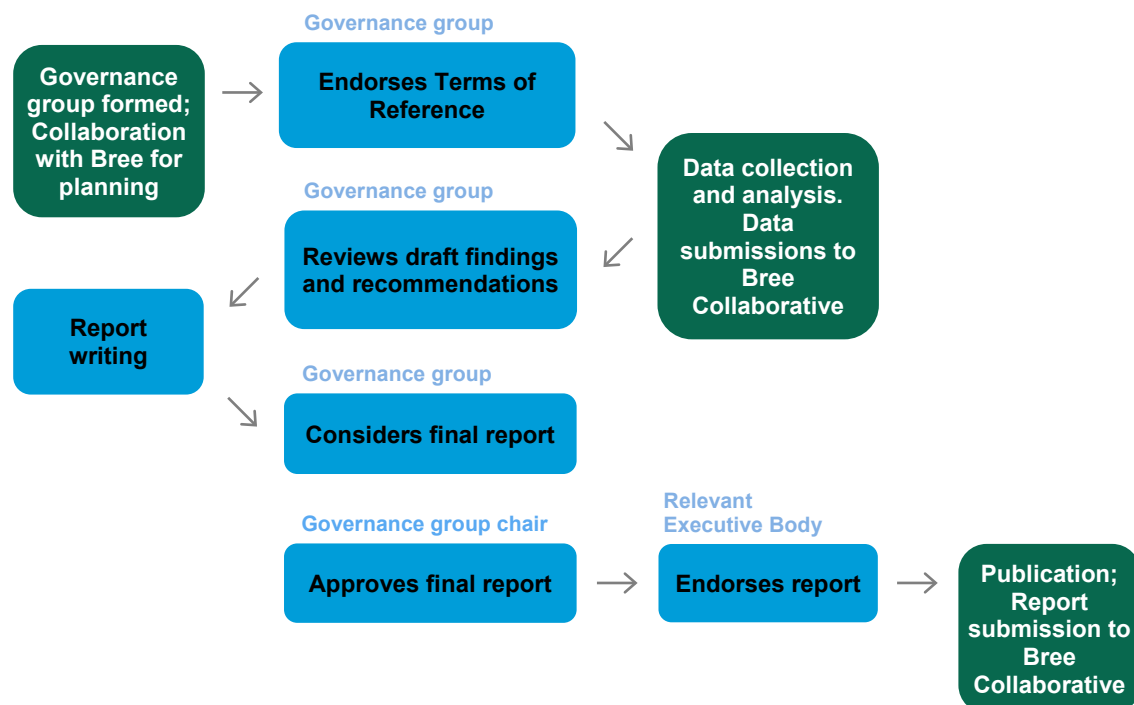


Table 4.2.1: Creating a timeline that considers other initiatives

Organizations using this framework should create a timeline for evaluation that considers alignment with Washington State HCA empower, National Health Strategy, FEMA community grants, Washington State Climate Resilience Strategy, and other federal and local heat, wildfire, and climate change initiatives and recommendations for other system-actors in the Bree Guidelines for *Health Impacts of Extreme Heat and Wildfire Smoke*, health systems may want to consider developing a timeline that considers new provider training in academic settings,

The timeline for organizational level evaluations should be detailed enough to help individuals external to the organization put the evaluation into a state-wide context.

Initiatives and programs	Start	End
Washington State HCA emPOWER improvements	TBD	
National Heat Strategy	2024	June 2030
FEMA community grants	N/A	
<u>Washington State Climate Resilience Strategy</u>	2024	

Timelines for evaluation should also consider the aims of the guidelines, which are to reduce impacts of extreme heat and wildfire smoke events on patients, and other organizational-

internal recommendations such as infrastructure development or patient education to make sure that all elements of the intervention have been at least partially or fully implemented before evaluating.

5.5 General Methods

A mix of methods, both quantitative and qualitative, should be used to gather evidence to answer the evaluation questions in order to provide a full picture of patient, staff, and other collaborators experiences, in addition to outcomes and impact data, depending on the type and number of evaluations each organization wishes to conduct. Methodologies should support, at least in part, an understanding of concordance of care with Bree recommendations and/or should aim to quantify the outcomes and impact of using the guidelines.

Methodologies for evaluations specific to each system actor should be agreed by the governance body prior to the commencement of each evaluation and should align with the guidance for each evaluation type listed in section 2.

Strong recommendations: Evaluations are expected to include in whole or part -

- Bree Collaborative Score Cards to support process or program evaluations;
- Desktop research: a systematic review of program documents which may include program guidelines, executed grant agreements, program logic, policy papers, and program reporting and procedure manuals. This may also include a review of relevant reports and existing data;
- Leveraging of other Foundation for Health Care Quality programs (e.g. Patient Safety), where applicable
- Data sampling, where applicable

Soft recommendations: Evaluations may include the following -

- Literature review: a systematic review of similar programs run in other jurisdictions, reviews or evaluations of similar programs, and relevant journal research articles or media reports (with caution)
- Semi-structured interviews with a range of stakeholders which may include face-to-face, telephone, or video-conferencing, etc.
- Surveys
- Economic profiling of the organization and region
- Case studies of selected projects or patient cases

5.6 Risks and limitations

When developing an evaluation[s] using this framework, organisations should consider the following risks and limitations as they pertain to demonstrating concordance of care, outcomes, or impacts associated with the implementation of the Bree Guidelines on OUD Treatment:

- Availability of resources and skills to conduct the evaluation/s

- Availability and quality of data from internal and external sources
- The burden/cost of collecting robust data
- Proportion of the program or initiative that can be directly contributed to the Bree Collaborative Guidelines and the difficulties or limitations of quantifying guidelines contributions
- Generalizability of the evaluation

These risk and limitations are ones that have been identified by the Bree as the primary one's pertaining to the ability to measure the outcomes and impacts of the implementation of these guidelines with reliability and fidelity.

The Bree Collaborative and the Foundation for Health Care Quality seek to mitigate some of these risks or limitations by offering resources for control of data collection limitations, data sharing limitations, and metrics and methodological alignment limitations that are found throughout this framework and in Bree and Foundation for Health Care Quality programs.

Table 4.4.1: Risks and controls [shade the 'Rating' cells as appropriate using the table below]

Risk	Results	Likelihood	Consequence	Rating	Control
Insufficient resources to undertake the evaluation	Low quality evaluation report; failure to meet timeframes; stakeholder dissatisfaction; damage to reputation of the organization	Likely	Fewer organizations are willing to conduct evaluations; effects of guidelines across the health care eco-system have gaps in knowledge	Substantial/High	Bree staff to consult on the evaluation design and methods; resources (templates, trainings, etc.) for implementation and evaluation planning; emPOWER map
Inadequate data to support analysis	Inadequate evidence to support findings; low quality evaluation report; stakeholder dissatisfaction; damage to reputation of organization	Possible	Understanding of guideline impact is reduced or incomplete	Substantial/High	Agreed evaluation matrix identifying objectives, goals, and metrics; data collection methodology (e.g. score cards); measures alignment (e.g. measurement bank)
Inability to untangle impacts of other initiatives	Lack of clear impact; diluted/exaggerated impact	Almost Certain	Inability to quantify the exact contribution of the Bree Collaborative work to system-wide changes	Minimal/Medium	Identification of common contextual factors; timeline alignment with other initiatives
Generalizability of evaluations	Fragmented evidence; evaluations irrelevant for state or nation-wide use	Possible	Inability to spread Bree best practices	Moderate/High	Survey question bank; evaluation framework;

The evaluation team should monitor the evaluation closely to ensure that these and other emerging risks are managed effectively. Table 2.4.2 defines the risk ratings used above.

Table 4.4.2: Risk ratings

Likelihood rating	Consequence rating				
	Insignificant	Minimal	Moderate	Substantial	Severe
Almost certain	Minor	Medium	High	Very high	Very high
Likely	Minor	Medium	Medium	High	Very High
Possible	Low	Minor	Medium	High	Very High
Unlikely	Low	Minor	Minor	Medium	High
Rare	Low	Low	Minor	Medium	High

Appendix A Theory of Change

The subcommittee for the Health Impacts of Extreme Heat and Wildfire Smoke report has not yet created a theory of change to accompany this report. Please refer to our Implementation Guide for updates.

Appendix B Data Collection Matrix

Evaluation Questions		Data: What to collect? When to collect it?			Data source: WHERE is it? HOW to collect it? WHO is responsible? ARE permissions required?
Questions	Indicators	Metrics	Context	Data Frequency	Recommended data source
Process/structural improvement					
What changes were made to patient identification policies or process?			<ul style="list-style-type: none"> Urban/semi-urban/rural Insurance type Population of focus 	Aligned with clinical considerations	Policies; workflows; QI programs; patient records;
What changes were made to the treatment initiation process?			<ul style="list-style-type: none"> Urban/semi-urban/rural Insurance type Population of focus 	Aligned with clinical and patient considerations	
What changes were made to policies or process for prescribing and continuation of pharmacotherapy?			<ul style="list-style-type: none"> Urban/semi-urban/rural Insurance type Population of focus 	Aligned with clinical and patient considerations	
What changes were made to clinician/patient/staff education?			<ul style="list-style-type: none"> Urban/semi-urban/rural Insurance type Population of focus 	Aligned with clinical and eco-system considerations	
What changes were made to patient access to services?			<ul style="list-style-type: none"> Urban/semi-urban/rural Insurance type Population of focus 	Aligned with clinical and patient considerations	
What changes were made to data sharing protocols or processes?			<ul style="list-style-type: none"> Urban/semi-urban/rural Insurance type Population of focus 	Aligned with clinical and eco-system considerations	
What changes were made to financial contracts or coverage policies?			<ul style="list-style-type: none"> Insurance type 	Aligned with clinical and eco-system considerations	
Effectiveness					
How effective were care coordination activities prevention?					

Outcomes					
EMS response for HRI/smoke exposure		See Section 2.1 – Capacity Measures			
Heat and smoke related ER visits		See Section 2.1 – Capacity Measures			
Impact of Guidelines					
Heat related mortality		See Section 2.1 – Mortality measures			(EMS data); APCD/HCA/DOH
Heat and smoke related morbidity (example: exacerbation of co-morbid conditions)		See Section 2.1 – Morbidity Measures			DOH/APCD/HCA
Other patient benefits? (economic, health, etc.)					
Lessons Learned					
Barriers and facilitators					Surveys; structured interviews; program documents;
“Pinch-points”					
	<p>What are you going to track?</p> <p>The concept that will help answer the question</p>	<p>How are you going to track it?</p> <p>How the concept will be measured</p>	<p>What will the indicators be compared to?</p> <p>For example:</p> <ul style="list-style-type: none"> specified target values baseline values a relevant benchmark or standard <p>a comparison group of comparable non-participants</p>	<p>How often will the indicators be collected?</p> <p>For example:</p> <ul style="list-style-type: none"> Weekly Monthly Quarterly <p>Annually</p>	<p>Program management team via program administrative data. This includes application forms, funding agreements, progress/completion reports, fees collected number of recipients etc.</p> <p>Policy team via program policy documents, media reports, etc.</p> <p>Evaluator via program documentation and/or literature reviews in collaboration with program/policy teams</p> <p>Evaluator via internal or external surveys or interviews and comparative data in collaboration with program/policy teams, data professionals, linked datasets or others as required</p>