

Evaluation Guide
New York City Health Care Coalition
Emergency Management Programs

New York City Department of Health & Mental Hygiene
Office of Emergency Preparedness and Response
Bureau of Healthcare and Community Readiness

A Guide to Evaluation for NYC HCC Programs

Contents

Questions about this guide	Error! Bookmark not defined.
Purpose of this document	4
What you'll find in this guide	4
Evaluation of Programs and Projects	4
Program Theory.....	4
Types of evaluation	5
Examples of impacts that could be found by evaluation	5
Examples of process evaluation	5
What's the difference between an activity, an outcome and an output?.....	5
Evaluating Health Equity	6
Helpful practices for equity data use	6
Writing goals and objectives	7
Goals vs objectives	Error! Bookmark not defined.
SMART Objectives Checklist	7
Writing for impact	8
Reporting Impact.....	8
Writing a data gathering plan	8
Additional tools for writing a data collection plan.....	10
General rules of thumb for gathering data	10
Increasing your survey response rate	13
Standard data to gather for all events	13
Evaluating training for preparedness and response	14
Application of Kirkpatrick's evaluation levels	15
Evaluating learning (Level 2) in emergency preparedness and response.....	15
Writing learning objectives	15
Pre- and Post-tests	16
Exercise evaluation methods	17
BHSR Program evaluation staff roles	17
APPENDICES	19
Appendix 1. Example 1 of Data Gathering Plan Table	19
Appendix 2. Example 2 of Data Gathering Plan Table.	21
Appendix 3. Developing a survey, step-by-step ("Survey Boot Camp").	22

A Guide to Evaluation for NYC HCC Programs

About this guide

The New York City Department of Health and Mental Hygiene (DOHMH) is providing this guide to a broad audience of its own Program Managers and those healthcare or emergency management organizations participating in the [ASPR/HPP](#). Some sections have direct applicability to DOHMH staff and others are applicable to all intended audiences.

Please direct questions about this guide to Darrin Pruitt at dpruitt@health.nyc.gov. For any links to resources found in this guide that do not work, report these to Darrin Pruitt at dpruitt@health.nyc.gov or ask your DOHMH Program Manager for assistance. To request any resource found in this guide as a Microsoft Word or editable PDF, please ask your DOHMH Program Manager.

Purpose of this document

This guide will help you with developing your own evaluation tools and/or with communicating with vendors about the tools they develop to gather data that assesses such things as training effectiveness (pre- and post-tests), audience feedback (post-event evaluation) and emergency preparedness exercises.

This guide is not intended to make you an evaluator. It is intended to provide you enough information to understand the activities and roles you play in evaluation as you structure and implement your projects and programs.

Darrin Pruitt offers review services for all the surveys and tests mentioned in this guide as well as exercise evaluation tools.

Darrin Pruitt
dpruitt@health.nyc.gov
347-396-2699

What you'll find in this guide

This guide provides an introduction to the following topics:

- | | |
|---|---------------------------------------|
| 1. Program Theory (or Change Theory) | 8. Evaluating training for readiness |
| 2. Types of evaluation | 9. Writing learning objectives |
| 3. Evaluating Health Equity | 10. Pre- and post-tests |
| 4. Writing goals and objectives | 11. Exercise evaluation |
| 5. Writing a data gathering plan | 12. Evaluating vendor product quality |
| 6. Rules of thumb for gathering data | 13. Roles and responsibilities |
| 7. Standard data to gather for all events and tools | 14. Appendices |

Evaluation of Programs and Projects

The Department of Health and Mental Hygiene (DOHMH) supports a comprehensive approach to evaluating the programs and projects it funds with money from the annual cooperative agreement from the [Administration for Strategic Preparedness and Response](#), the [Hospital Preparedness Program](#) (HPP). DOHMH extends this approach often to programs and projects that are not funded. DOHMH evaluation staff work with program managers and HPP sub-recipients and other contractors to develop approaches to evaluation and reporting each year. We select up to two subrecipients yearly for “making sense” evaluation; that is, we avoid making judgments of performance. We instead look to causes for underperformance and ways to address those causes. DOHMH is also on the “receiving end” of such evaluations, looking to the subrecipient being evaluated for input on how DOHMH can do a better job of implementing the HPP, develop programming to support the subrecipient or their sector and make overall improvements to contracting and the program overall. DOHMH also reviews Healthcare Preparedness Program deliverables and the NYC Health Care Coalition activities yearly, culminating in 5-year retrospective and forward planning document.

Program Theory

A program theory is a big picture, strategy-level map that shows activities and interventions at work in bringing about desired change. It shows all pathways to the change of interest and provides context for the change process. Ideally, it should show progression of a problem to an aspirational or realized impact (or change). Program theories are normally created in partnership between evaluators and stakeholders to improve accuracy and buy-in. Developing a program theory is important so that you know how your resources and planned activities are acting or interacting to bring about change. You can test your theory once it's

A Guide to Evaluation for NYC HCC Programs

established and remove or add resources or interventions that are not working to bring about change or are not doing so effectively and efficiently. A deeper explanation and examples of program theories can be found [here](#) and [here](#). Also see *Developing Monitoring and Evaluation Frameworks, 1st Edition* by Anne Markiewicz and Ian Patrick.

Types of evaluation

Impact evaluation is taking into account the measures taken or known forces during a program to assess how they affect or change the target subject, population, sector, etc. These effects or changes may be positive, negative, intended or unintended. Usually, the intended impact of a program is expressed before program implementation in the form of goals and objectives.

Examples of impacts that could be found by evaluation

- Counting the number or increase in the number of requests for information a nursing home can answer due to training on situational awareness software;
- Demonstrating the ability of a community health center to continue operations for the patients it normally sees in a new location or mobile medical unit during a disaster; and
- Demonstrating correct triage and isolation of a patient with a communicable disease of public health concern due to successful completion of mystery patient drills.

Process evaluation is the assessment of how an organization implemented a program, the steps to be followed and if those steps of implementation affected the program's impact. Process evaluation is also helpful for improving processes (efficiency, reach of program, user satisfaction, etc.) Questions asked for process evaluation include "Are there any identified gaps or barriers to obtaining our intended outcome?" or "What can be done to address barriers to meeting the program goals?"

Examples of process evaluation

- Determining how many webinars planned for the long-term care program were actually completed;
- Describing why a workshop worked as planned and another did not;
- Determining if parts of an emergency response plan work as they are currently written and how to address those that do not;
- Determining why some meetings have better attendance than others; and
- Pre- and post-tests for training can be considered part of a program's process evaluation. (See below.)

What's the difference between an activity, an outcome and an output?

It's helpful to get an understanding of the differences in these terms as they are often used somewhat interchangeably when they are very distinctly different from one another.

It is easy to confuse an activity you undertake in your program for an output. To see the difference, ask yourself when writing activities what they are accomplishing or producing.

A Guide to Evaluation for NYC HCC Programs

Activities vs. Outputs: Outputs are the direct, immediate-term results associated with an activity. In other words, they are usually what the activity has produced.

Activities: “What we do...”	Outputs: “What we produce...”
Develop plans and protocols	All-Hazards communications protocol with EEs
Provide training for partners	Train the trainer guide
Attend HCC quarterly meetings	Certifications or credentials earned

Outputs vs. Outcomes: Outcomes are the change that occurs as a result of an activity.

Outputs: “What we produced...”	Outcomes: “What changed...”
All-Hazards communications protocol with EEs	Faster assessment of needs during a response
Train the trainer guide	Lower cost training, higher retention of skills
Certifications or credentials earned	Credentialed staff can lead response activities.

Evaluating Health Equity

The City of New York is committed to advancing racial equity, and DOHMH specifically is dedicated to promoting health equity to achieve fair health outcomes for all New Yorkers. OEPR for Equity developed a tool to help program developers and managers ensure the principles of racial and social justice are included in their programs. BHCR staff should consider utilizing this tool as a resource for their projects. Subrecipients and vendors with contracts managed by DOHMH/BHCR staff should consult with their program manager on incorporating racial equity and social justice into their contract deliverables. You may also consult the [New York City Race to Justice web site](#).

Please contact the [O4E coordinator](#), Tala Mansi, if you have any questions on the tool and its implementation.

Helpful practices for equity data use

An evaluation process can help determine whether programs and services are reaching and meeting the needs of diverse racial, ethnic, and Limited English Proficiency populations. Collect your own data to compare the service population to the populations actually using or participating in the program to determine whether the provision of services discriminates against certain populations and address any such gaps in services.

Use data when planning for emergency response, as well as for long term recovery and mitigation efforts. Census data or data collected by DHS, HUD, FEMA, school systems, community organizations, state and local government agencies, and third parties can be used to ensure evacuation activities and other emergency response efforts are developed that best serve the communities in the service area. Use resources such as the Census Bureau and the Census Information Centers for training and technical assistance on

A Guide to Evaluation for NYC HCC Programs

effective demographic analysis for targeted outreach. Greater New York Hospital Association (GNYHA) also has a page of data resources for their members. Among them are equity data use guidelines.

<https://www.census.gov/about/what/data-equity/tools.html>

www.lep.gov

[GNYHA Demographic Data Collection](#)

Data can include the geographic distribution of individuals by race, color, and national origin (including limited English proficiency); the reliance of particular communities on public transit; the proximity of different groups to emergency or disaster danger zones or factors affecting their health; and the geographic distribution of damage, taking into account varying degrees of severity.

Writing goals and objectives

Generally speaking, all BHCR Program Managers are aiming at the same goal: To give all facility types in all healthcare sectors the capability and capacity to prepare for, respond to and recover from emergencies affecting the healthcare system in NYC.

How we get to that goal is expressed in objectives, which are more specific and focused.

Evaluation is almost completely dependent on writing clear goals and objectives.

SMARTIE objectives are particularly essential to measuring both process and impact of programs. Click [here](#) to see more information from the CDC.

TIP: Write your SMARTIE objectives with input from your emergency management team, community advisory boards and DOHMH Program Manager.

SMARTIE Objectives Checklist

- ☐ **Specific:** who and what are clear.
- ☐ **Measurable:** how much, how many, when, how quickly, etc. are clear.
- ☐ **Achievable:** the “what” can be done by the “who” within the timeframe listed.
- ☐ **Realistic:** resources are available and/or constraints can be managed.
- ☐ **Time-referenced:** when or by when is clear.
- ☐ **Inclusive:** representation of the racially or socially marginalized is considered.
- ☐ **Equitable:** how social justice or health equity is addressed is included.

Writing for impact

We set goals and report on them for ourselves and developing our programs. We also do this for those interested in the change or impact due to our efforts and our return on the considerable financial investment of HPP and [PHEP](#) funds. For this reason, we develop a written impact report annually.

Impact is the reportable, quantifiable, or potential difference that your program is making in people's lives. We seek to report payoffs and benefits to NYC's healthcare system or healthcare coalition. The focus is on the New Yorker's benefit – not internal, organizational or personal benefit.

Reporting Impact

The evaluation unit of BHCR can help you develop statements of intended impact for your programs. This would be a key first step for program development. In these statements, describe the change you will have on preparedness, response and recovery capacity for your healthcare sector. In the spring each year, you can use these impact statements to plan for deliverables in the new budget period to come. The changes you and your vendors accomplished through deliverables is the **impact** your work has had.

See [tips for impact writing](#).

Writing a data gathering plan

After you have written your SMART objectives, use them to develop your plan steps and tactics. Most plans will yield a great many data points, either observable and quantifiable outputs and outcomes that you can track, or narratives about the process or changes the plan has made. As you roll out your plan, this is the time to plan for what data you wish to gather and how. Developing a data gathering plan entails most or all of the following steps:

1. Review of SMARTIE objectives

Pick out the data points that will support your observation of changes. That is, select things you can observe such as additional members, increases in the number of completed plans, or increases in conducted exercises in which sectors participate.

2. Select meaningful data points

Data points can include numbers, dates, averages, narratives, etc. that answer:

- Time? How much quicker than last year? How much slower?
- Number? How many? How *many more* than last year? How many more than when we started the program? Number achieved versus number required (or planned)? How many of a certain type?
- What score? What score for group x versus group y? How much are scores changing for x group or staying the same?
- All? Partial? None?
- Specific names, models, brands?
- Why? How? When?
- How have participants' attitudes towards X changed? How have their behaviors in X situation changed?
- Is the quality of the program improved by the program's activities? Was the content of the event more relevant than that of the last event, or the content presented last year?

3. Ensure you are able to collect each data point

Decide for each data point:

- Can be captured along with others or needs to be captured separately?

A Guide to Evaluation for NYC HCC Programs

- In what venue (e.g., meeting, webinar, training) and in what conditions (e.g., when everyone is there, when the nursing homes have completed 2 learning sessions) will it be best to capture the data?
- How many times will it need to be captured (one time or multiple times to discover trends)?
- What means (e.g., survey, pre- and post-test, observation with a checklist) or what tools (e.g., tablet, audience participation devices) will work best to capture it?
- Who can capture the data? Do they need training to do so properly (without bias or errors)?
- Is it best to have more than one evaluator or observer capture the data at the same time? Things that are complicated to observe benefit from more than one evaluator. Subjectivity or bias may be mitigated by having more than one observer and comparing or averaging their scores.

4. Plan for the kind of analysis you'll need for each data point or data points

- You may only wish to get counts, totals, averages (means), modes or medians. This will describe the outcomes. These are [descriptive statistics](#).
- If you want deeper analyses that allow you to compare, assess the strength of associations between data points, predict or model the effect of one data point on another or describe cause and effect—that is, *make inferences*—you'll need to use [inferential statistics](#). These analyses are more complex and usually require careful consideration of several other factors (assumptions) in order to use them as they are intended. You can read more about inferential statistics [here](#).
- You may also want to ask open-ended questions to obtain qualitative or narrative data that provides context and additional information to descriptive or inferential statistics.
- Further, qualitative methods may be helpful for setting up surveys or vice versa. For instance, focus group findings could help determine which survey questions to ask, or answers to survey questions can further inform focus group questions.

5. Organize your list of data points

Put your data points in a table, organized by the program objectives. Determine in columns the best way to obtain the data, point my point. This should allow you to pinpoint the source of the data and the method for gathering it. See [Appendices 1](#) and [2](#) for examples.

6. Group your data points

Group your data points together according to commonalities (e.g., they can be gathered at the same time by the same person(s) at similar events).

7. Write the instruments

Write the surveys, tests and/or observation forms (e.g., EEGs for exercises, checklists for observers) you intend to use.

8. Train your evaluators and observers to use the instruments.

9. Plan for a pilot of your instruments (see below).

10. Carry out the analysis

Once piloted and data gathered, carry out the analysis of the data.

Additional tools for writing a data collection plan

To get a reliable set of data that answers your questions, you need to ensure you are asking the right questions. Additional help can be found using the links below.

- [CDCs Data Collection Methods for Program Evaluation](#)
- [NYC Health Department Data Page](#)
- [Data Collection | Definition, Methods & Examples \(scribbr.com\)](#)
- [CDCs Evaluation Tools and Templates](#) (including Emergency Response)

General rules of thumb for gathering data

1. Follow the plan

Follow the plan you have made addressing how to get the data for your SMART program objectives.

2. Develop surveys well ahead of time

Develop your surveys and/or pre- and post-tests well ahead of your events or the time you intend to launch your data gathering efforts. This will give you time to consult with the BHCR evaluation and data team and make changes.

3. Pilot your survey

If at all possible, pilot your survey to see if there are technical problems or if your respondents will have trouble understanding the questions. A best practice is to pilot the survey questions with a similar population or group that you will be targeting. If that is not possible, try piloting with someone less familiar with the project to make sure that the questions will make sense to a person who hasn't been involved in the program's development.

4. Align pre and post-test questions with learning objectives

Align pre and post-test questions with learning objectives for training established ahead of developing the training materials. (See below for more on developing pre and post-tests.)

5. Align items on surveys with program objectives

All items on surveys should provide data to the program manager that supports reporting that a program objective has been met, whether partially or completely. Items can also provide supporting data such as time spent on technical assistance, difficulties making time for webinars or challenges to engaging stakeholders. Those items that do not provide supporting data for objectives either need to be rewritten or indicate the program objectives themselves may be incomplete.

6. Ensure survey items ask for data that is **observable and actionable**

Ensure all items on surveys and other instruments gather observable and usable data. For example, questions that ask program participants if they liked the program overall do not provide actionable data. Questions that ask if objectives have been met (and list the objectives) can tell a program manager where to target corrections.

A Guide to Evaluation for NYC HCC Programs

7. Ensure survey items are clear and easily understood by your audience

Essential to getting good data is doing everything possible to ensure your respondents understand what you are asking.

- a. Take steps to ensure your survey is going to the **audience who understands the topics** and can answer your questions.
- b. **Explain why you are asking your respondents to answer your survey** in a brief introduction.
- c. **Arrange items on a survey by topic.** It is often easier for the respondent to answer questions they see as referring to a specific thing. A good practice is to group the questions on a topic together under a heading with a short sentence or two about the context of those items. Respondents may have received your program in a specific order of topics. Use that same order. Respondents who work in specific professions often think of certain things as going together. For example, nurses may think of screening and triage together; so questions about screening and triage need to be together. If you don't know the order of topics or how topics are perceived as going together, this can be addressed in your pilot of the survey.
- d. **Avoid vagueness.** If you ask a respondent if a program has been "effective," define "effective." Better still ask specifically, for example, if a certain aspect of their programs, procedures, knowledge, etc. has changed and how it has changed. An example of a vague question is "Was our module on developing an Emergency Operations Plan effective?" A stronger question is "How confident are you that you can develop an Emergency Operations Plan, now that you have completed the module on developing one?" Responses could be "I am highly confident that I can do this now on my own," "I am moderately confident that I can do this now on my own," "I have little confidence that I can do this now on my own," and "I have no confidence that I can do this now on my own." You could also add this same list of items but change "on my own" to "with assistance from an expert or consultant."
- e. **Make sure choices for survey items are distinct from one another.** Items like this are said to be *mutually exclusive*. That is, there is no overlap in meaning between items. For example, in this question the choices are not distinct from one another:
 - i. "How would you rate your ability to develop exercise objectives for your evacuation plan?"
 1. Really high
 2. Great
 3. Moderate
 4. Semi-confident
 5. Not good

But in this question, the choices are clearly different from one another:

- ii. "How would you rate your ability to develop exercise objectives for your evacuation plan?"
 1. Very high
 2. High
 3. Moderate
 4. Low
 5. No ability

A Guide to Evaluation for NYC HCC Programs

- f. **Ensure survey items ask for responses about one thing and not two or more things.** In the next example, the respondent could be rating clarity or relevance. These two concepts are distinct from one another but included in the same survey item.

- i. “How would you rate the clarity and relevance of the program’s objectives?”

How to fix: Write this as two separate items or decide which is more important to you (or matches your program objectives) and just write one.

- ii. An exception is when two or more things are always or almost always done together or are thought of being done together. Examples:

1. “To what degree are you aware of your organization’s policies on *diversity, equity and inclusion*?”
2. “Rate your understanding of *decontamination and waste management* procedures.”
3. “How well did this functional exercise test *command and control* issues for your healthcare network?”

- g. **Include demographic items.** Make sure to include demographic items or items that can distinguish or group respondents (e.g., hospital name, facility type, etc.). This allows you to describe subgroups of respondents in the analysis. Limit the demographic questions to only those that are important for your analysis. See “[Standard Data to Gather for All Events and Tools](#)” for examples.

8. Ensure data can be tracked over time

Use similar questions for surveys or pre- and post-tests across years of a program, where applicable, so program needs and improvements can be observed over time.

A Guide to Evaluation for NYC HCC Programs

Increasing your survey response rate

To get a reliable set of data that answers your questions, you need the highest response rate possible. To ensure a high response rate, particularly among respondents who have little time and who are bombarded by surveys, here are some suggestions:

- At least 4 weeks prior to sending your survey, determine who your point person on at DOHMH will be. This is most often the program manager with whom you work on contract deliverables. That person can help you determine how to
 1. send out the survey,
 2. answer questions about items,
 3. develop FAQs and post them (or send them to the respondent point(s) of contact) and
 4. monitor the response rate.
- Determine who your points of contact are among the respondents (facility, network, clinic or organization). Get a clear and correct list of their contact information.
- At least 2 weeks prior to when you plan to send the survey, send these points of contact a message describing
 1. the survey purpose,
 2. how the survey relates to preparedness response or recovery,
 3. what you plan to do with the results
 4. who needs to respond,
 5. how respondents' information and responses are protected,
 6. how long it will likely take to complete it,
 7. when they will get the survey and
 8. when it is due.

You may ask for a list of respondents' emails or if the point of contact wishes to distribute (and monitor) the survey themselves. It is a good practice to include a PDF version of the survey in this communication with your points of contact and stress it should be used to gather the best group of respondents together to answer the survey. The PDF can also be used to see the progression of questions and plan for the responses to each.

- One week prior to launching the survey, test your survey on handheld devices, such as mobile phones or tablets, to leverage the flexibility respondents may wish to have in answering when they cannot access a computer.
- On the day of the survey launch, send the survey. Send **both** the link and the PDF so respondents can share the PDF and get answers from colleagues and so they can see all the questions at once. (Often online survey tools are paginated and prevent respondents from getting a sense of the whole survey.) Send the same information as above in 7b and c.
- A week before the survey is due check the response rate and communicate with your point of contact about getting more respondents to complete the survey.
- After the due date, permitted this is possible, ask your point of contact to see if more respondents can answer.

Standard data to gather for all events

To facilitate comparing data across events or surveys from year to year, BHCR program managers and vendors can use standard data on all their data-gathering instruments. The below lists some practices for doing this. The items are in an ideal order -- as they should appear in such tools as pre- and post-tests and session evaluation surveys -- that will also facilitate comparing data sets and performing analyses.

Table 1. Standard data points and best practices for gathering them for all BHCR-related events and data gathering tools.	
Survey item	Format
Date	MM/DD/YYYY
Name (if not anonymous)	Last Name, First Name
Organization or Facility Name Facility*	Provide a drop down so that the name is always used the same way by all BHCR programs
Sector	Provide a drop down of "acute care," "ambulatory care (primary care/community health)," "ambulatory care (dialysis)," "ambulatory care

A Guide to Evaluation for NYC HCC Programs

Table 1. Standard data points and best practices for gathering them for all BHCR-related events and data gathering tools.	
Survey item	Format
	(urgent care),” “long term care (nursing home),” “long term care (adult care facility),” “long term care (home care),” “long term care (hospice and palliative care)” and “other”.
Event Type	Provide a drop down of “training,” “exercise (tabletop),” “exercise (drill),” “exercise (functional),” “exercise (full scale),” and “meeting or conference” or include in the survey name).
Event Name	Include event name in the survey name
Rating scales for surveys asking for a reaction or perception of value**	Provide a single choice list of the following responses, called anchors: “excellent,” “good,” “fair,” “poor,” and “very poor”

*Names are kept by the healthcare coalition planning and programs unit for primary care centers (community health centers, both Federally Qualified Health Centers (FQHCs) and non-FQHCs), dialysis centers, opioid treatment programs, nursing homes and adult care facilities.

**You may switch scales – both in the order of responses and the anchors for them – but do so with care. It is a good idea to separate sections of your survey and provide a few lines of introduction when you intend to change the scale. Include language in the introduction to the new section to alert the respondent to the change and what you intend to understand from their responses.

Evaluating training for readiness

While this guide focuses on evaluation and data collection to support it, it is helpful to provide some background on training and its place in improving capabilities for emergency preparedness and response. Many of BHCR’s program activities involve training of health and allied health professionals with the aim to change their knowledge about plans and procedures as well as their skills to execute their tasks or play their roles in those plans and procedures. Training evaluation is only valid for well-designed training with well-designed pre- and post-tests and follow up evaluation. Below are some of the considerations you should make in designing training that supports your mission and can stand up to evaluation.

1. Design training following the Analysis > Design > Development > Implementation > Evaluation (ADDIE) model or similar model. See the [ADDIE model](#).
2. If each of the ADDIE phases is executed thoroughly, the various stages of evaluation can tell the program manager and vendor about the changes brought about by the training.
3. Evaluation of training is most often thought of as the pre- and post-test. But other levels of evaluation play a role in understanding the change brought about by training. Most often [Kirkpatrick’s framework](#) for evaluation of training, including four levels of evaluation, is used. See below.

A Guide to Evaluation for NYC HCC Programs

Application of Kirkpatrick's evaluation levels

Level 1. This is most often accomplished with a “reaction survey,” asking trainees about the venue, the quality of the materials, the preparedness of the trainer or the relevance of the training to their roles or jobs. Problems here can prevent change in knowledge and skills, the focus of level 2.

Level 2. This evaluation is most often accomplished via pre- and post-tests. It is important to note that level 2 evaluation is seeking to record change in knowledge and skills. Knowledge can be tested on well-developed pre- and post-tests. *Skills should be demonstrated by the trainee, observed and scored.* Subject matter experts should design pre- and post-tests as well as observation forms. Subject matter experts should serve as observers for trainee demonstration of skills. If there are no or less than expected changes in knowledge and skills, training content or delivery methods most likely need to be revised and training needs to take place again. Level 3 depends in part on seeing changes at level 2.

Level 3. Most often BHCR program managers and vendors cannot observe changes in behavior or the application of knowledge and skills in the work environment needed for this level of evaluation. However, program managers and vendors can do follow up surveys about these changes by asking the trainees themselves or their supervisors. Note that failure to perform at this level or make changes is not necessarily a reflection of the quality of the training. The work environment can have an impact by posing barriers to change or the trainees can fail to perform because their duties or regulations governing their duties changed.

Level 4. This evaluation looks the organization or system's bottom line, the reason training or changes in performance were needed in the first place. For preparedness and response, this is most likely driven by the documented or perceived need to train certain staff or staff in certain sectors so they can coordinate with those staff who already know or have the skills to work to prepare for or respond to emergencies. This type of evaluation is best accomplished through exercises or real events and is the purpose of after-action reports (AARs).

Evaluating learning (Level 2) in emergency preparedness and response

[The Public Health Foundation](#) offers this [comprehensive guide](#) to assess learning (Level 2 evaluation) for each program trainings you deliver. The tables you find within this guide list learning objectives by domains 1 to 4.

- Domain 1 is “model leadership.”
- Domain 2 is “communicate and manage information.”
- Domain 3 is “plan for and improve practice.”
- Domain 4 is “protect worker health and safety.”

[The PERLC Network](#) has other robust resources you may find helpful for evaluating training programs in emergency preparedness and response.

Writing learning objectives

Writing learning objectives is at the heart of being able to evaluate training. Clearly written learning objectives allow for writing good questions in pre- and post-tests. Questions should be written to the learning objectives and if they are, results of the tests will indicate to you whether or not training brought about changes in knowledge and skills among the trainees. Like program SMARTIE objectives, learning objectives are observable and thus measurable through either a pre- and post-test or an observation sheet for observing trainees performing tasks. A couple of sites can provide you an ample tutorial and resources for writing learning objectives. See [writing learning objectives](#) and the [objectives builder application](#) online.

Pre- and Post-tests

Pre- and post-tests can be considered both process and impact evaluation. “Process” because they describe the outcome of training (what is often an initial or foundational step in many programs). “Impact” because they describe a change to knowledge and skills. Mostly likely, however, they provide only a part of the data needed to assess program impact because the change in knowledge and skills lays an essential foundation for changes in behavior, or the use of knowledge and skills, but *does not ensure those changes occur in the work environment*. Nonetheless, these tests should be taken seriously.

From the pre- and post-test scores a program manager may see clues for changing training to make sure it provides content in clearer ways for trainees to obtain the knowledge they need. The post-test scores can also reveal which content, however clearly presented, is the most difficult for trainees to understand.

The purpose of pre- and post-tests is to evaluate an individual’s understanding and *change in understanding* due to training or some other educational intervention. In order to make these tests more effective in their purpose, here are some helpful suggestions.

1. Pre- and post-tests should follow the learning objectives set for the training or event. (Refer to the section above.)
2. It is a good practice to write at least two test questions for each learning objective. One will be used for the pre-test and the other for the post-test. Or you can use both in the pre- and post-tests and provide alternate versions of the tests so trainees cannot collaborate with one another.
3. Pre-tests should pose questions that truly discern those trainees who do not understand or have prior knowledge of a topic from those who do. Post-tests should pose questions that truly discern those trainees who have gained understanding or can apply knowledge of a topic due to the training from those who have not or cannot.
4. In order to accomplish the aims of pre- and post-tests, test items (questions) cannot have obvious right or wrong answers or be vulnerable to guessing. See guidance on writing good test items ([general](#), [another general guide](#) and specifically a [guide for multiple choice questions](#)).
5. In order to make a comparison between pre- and post-test scores and truly see the effect of the training on knowledge and skills trainees have gained, observe the following:
 - a. Ask the trainee to include his/her/their name on pre- and post-tests. If you wish to maintain anonymity, you may pre-fill tests with unique identifiers or ask trainees to use a unique identifier such as date and month of birth, their mother’s initials, etc. on both pre- and post-tests.
 - b. Record the sector (acute, long term, ambulatory (and sub type) on the test.
 - c. Record the trainee’s organization on the test (optional but may help if trainees have similar names).
 - d. Ensure trainees work on their own and do not collaborate or use online sources for their answers.
 - e. Administer and collect the pre-test first thing before starting the presentation of the content.
 - f. Use similar questions across years of a program, where applicable, so that program needs and improvements can be observed over time.

Exercise evaluation methods

Exercises are evaluations. They are intended to evaluate plans and procedures at an organizational or system level. They are not intended to evaluate individual performance. The process for developing exercises should follow the HSEEP guidelines. Click [here](#) to read more. You may also benefit from searching for exercise materials on [TRACIE](#), or take [FEMA Independent Study](#) courses online such as [IS-120](#), or [IS-130](#). In order to get the most out of the evaluation the exercise is intended to provide, follow these steps:

1. Consult with exercise and/or evaluation experts early in the exercise development process to be sure that the Exercise Evaluation Guide (EEG, the principal tool for exercise evaluation) will provide you the information you need to assess whether the objectives were met. This required information is generally about the performance of the organizations playing in the exercise against the plans and specified exercise objectives. You may also wish to observe other performance or conditions related to the objectives or having potential impact on performance.
2. List the plans the exercise is intended to test. Consult guides (such as [this one](#)) on writing exercise objectives.
3. Consult with your exercise design team in your concepts and objectives meeting or initial planning meeting to develop a clear, short set of exercise objectives related to the plans in step 1.
4. Develop the scenario and Master Scenario Events List (MSEL—for more complex, primarily operations-based exercises) in order to test these objectives.
5. Develop the EEG(s) so that evaluators can observe whether the players' responses to the scenario and injects align with the expected outcomes outlined in the plan or protocol you are testing. EEGs are not easy to develop. Here are some helpful links:
 - a. [Comprehensives slide set on exercise evaluation](#)
 - b. [Developing Exercise Evaluation Guides](#)
6. Provide points of clarification or an overall picture of what should happen during the exercise to your evaluators when you train them prior to the exercise. An exercise walk through is recommended, especially for full scale exercises with action taking place in one or more areas of play. See [sample controller and evaluator training](#).
7. Hold a hot wash post-exercise and provide player feedback forms to help you discover barriers to performing actions according to the plan and objectives.
8. Use both EEGs, note takers (if applicable), hot wash notes and player feedback forms in order to analyze what took place for each objective.

BHSR Program evaluation staff roles

Role: serve as lead in developing the evaluation plan, methods and tools to ensure the impact of BHCR's programs is clearly articulated for reporting and future planning, and yearly required reporting to ASPR HPP is completed fully and on time.

A Guide to Evaluation for NYC HCC Programs

1. **Develop and maintain an evaluation plan for BHCR's programs.** This may include targeting specific deliverables or sets of deliverables for evaluation and/or monitor them or other deliverables in order to complete HPP requirements for the budget period.
2. **Alert all Program Managers if there are deliverables in their contracts supporting program evaluation or HPP requirements.**
3. **Review all scopes of work outcome statements.** These statements can be incorporated into each deliverable. They should describe clearly what the healthcare system or particular sector (e.g., long term care) will have or the change the system will experience.
4. **Assist Program Managers in developing quick sheets for their sector** that include a brief overview of sector impacts already accomplished, impact level goals for the sector over the next 3 to 5 years and the impact intended by current budget period scopes of work.
5. **Develop impact statements from work completed by sub-recipients and contractors via their scopes of work** during the middle to late summer yearly in order to complete the HPP/BHCR annual report.
6. **Create opportunities for feedback from BHCR Program Managers on evaluation methods and incorporate this feedback** in a continuous cycle to streamline processes and increase accuracy and timeliness of measurement and reporting.
7. **Assist Program Managers with adhering to best practices for evaluation** by reviewing all data gathering tools including surveys, pre- and post-tests and EEGs and other exercise evaluation tools. Evaluation staff need at least one week's notice in order to carry out a thorough review.

APPENDICES

Appendix 1. Example 1 of Data Gathering Plan Table

The below is the data analysis plan for analysis of supply chain data received from NYC hospitals and networks via the supply chain survey, 2017. As this is for one survey, the data plan is rich but reasonably straightforward.

Research Question	Indicator	Sort Data By	Additional Steps Req'd?	Data Output
Network/Hospital Information				
1. How many acute care and non-acute care facilities are covered by this assessment?	Number of acute care facilities considered for this assessment Number of non-acute care facilities considered for this assessment	<ul style="list-style-type: none"> Independents vs networks Aggregate 	None	Tables (3) Or simple text description
2. What types of non-acute care facilities are covered by this assessment?	Free text next to number of non-acute care facilities considered for this assessment	<ul style="list-style-type: none"> Independents vs networks Aggregate 	Coding facilities by type	Frequency Table (3)
3. What activities will facilities undertake in the next ~2 years to develop situational awareness between emergency management and supply chain functions?	List activities your facility or network can undertake in the next ~2 years to develop situational awareness between emergency management and supply chain functions	<ul style="list-style-type: none"> Independents vs networks Aggregate Comparison 	Coding free response text	Frequency Table (if applicable) Simple text description
Methods for Assessing Supply Needs				
4. How are specific supply amounts estimated overall for networks and independent hospitals? Are they very different from each other?	How are specific supply amounts estimated overall?	<ul style="list-style-type: none"> By supply for independents vs networks Aggregate Comparison 	Coding free response text	Frequency Table (if applicable) Simple text description
5. How are specific supply amounts estimated for the 3-4 disaster scenarios listed in the assessment for networks and independent hospitals? Are they very different from each other?	For the 3 to 4 disaster scenarios listed below, how are specific supply amounts estimated?	<ul style="list-style-type: none"> By disaster scenario for independents vs networks Aggregate 	Coding free response text	Text description
Research Question	Indicator	Sort Data By	Additional Steps Req'd?	Data Output
Supply Needs, Operations and Delivery				
6. What are the top 3 critical resources facilities will need to maintain acute and non-acute facility operations?	What are the top 3 critical resources (materials or equipment) you will need to maintain acute and non-acute facility operations related to this	<ul style="list-style-type: none"> By disaster scenario for independents vs networks Aggregate 	Coding "Other" responses	Frequency Tables
7. Do facilities get all or part of their supplies and materials just in time?	Do you get all or some part of this material or equipment just in time?	<ul style="list-style-type: none"> By material for independents vs networks Aggregate 	None	Graphs (3) And/or List by material

A Guide to Evaluation for NYC HCC Programs

8. Do facilities normally store some amount of these materials for disasters?	Do you normally store some amount of this material for disasters?	<ul style="list-style-type: none"> By material for independents vs networks Aggregate 	None	Graphs (3) Or Lists by material (3)
9. How much of these materials do facilities think are available within their network or from a facility which they have a relationship with?	How much of this material do you think is available within your network or from a facility with which you have an agreement?	<ul style="list-style-type: none"> By material for independents vs networks Aggregate 	None	Graphs (3) Or Tables (3)
10. Can facilities use alternatives or equivalents for these materials?	Can you use an alternative or equivalent for this item?	<ul style="list-style-type: none"> By material for independents vs networks Aggregate 	None	Graphs (3) Or Tables
11. What problems have suppliers encountered in the past with deliveries during emergencies? Are there any discernable patterns in problems suppliers encountered by facility type?	What problems have your suppliers encountered in the past with deliveries during emergencies?	<ul style="list-style-type: none"> By material for independents vs networks Aggregate 	Summarize free response text; group by similar responses	Frequency tables (if applicable) Or Simple text description
Research Question	Indicator	Sort Data By	Additional Steps Req'd?	Data Output
12. Which materials or equipment procurement shifts according to season?	Please name materials or equipment whose procurement challenges shift according to season.	<ul style="list-style-type: none"> By material and season for independents vs network Aggregate Top 3-5 materials (aggregate) 	None	Tables
Suppliers/Distributors				
13. What are the names and locations of facilities' primary distribution centers for supplies? Do these distribution centers have emergency management or continuity of operations plans?	What is the name and location of your primary distribution centers for supplies? Do these distribution centers have emergency management or continuity of operations plans?	<ul style="list-style-type: none"> Independents vs networks By borough By distribution center 	<ul style="list-style-type: none"> Break out addresses by street, city, state, and zip code Add borough to addresses 	List
14. How many and which facilities are using the same distribution centers as their primary distribution centers for supplies?	Facility name and location What is the name and location of your primary distribution centers for supplies?	<ul style="list-style-type: none"> Independent vs networks Aggregate 	None	List
15. Who are the primary and secondary suppliers for these materials and resources?	Who is your primary supplier or distributor for these materials? Who are your secondary or other suppliers for these materials?	<ul style="list-style-type: none"> By supply/material Independents vs networks By borough 	None	Tables And Text description
Research Question	Indicator	Sort Data By	Additional Steps Req'd?	Data Output
16. Do the vendors have emergency management	Do your vendors have emergency management	<ul style="list-style-type: none"> By vendor 	None	Tables

A Guide to Evaluation for NYC HCC Programs

or continuity of operations plans?	or continuity of operations plans?			
NYC Healthcare Coalition Planning				
17. Which items would facilities like the NYC HCC leadership to prioritize?	List any items to the right for which you would like the NYC Healthcare Coalition leadership to prioritize policies or operational activities (e.g., situational reports)	<ul style="list-style-type: none"> • By rank • Independents vs networks 	None	Tables

Appendix 2. Example 2 of Data Gathering Plan Table.

This this example is for an entire program, supported by multiple data gathering methods, the table for the plan is large and complex. See: S:\EPR\BHPP\HSRE\Projects\Evaluation\Unit Folders\CDPP\Data Gathering

File name: CDPP data sources and methods, Tab: data view

For sub-recipients desiring to see this plan, please email dpriitt@health.nyc.gov.

A Guide to Evaluation for NYC HCC Programs

Appendix 3. Developing a survey, step-by-step (“Survey Boot Camp”).

Here are suggested steps to follow for developing a survey. Don’t be overwhelmed by this. This only meant to help you not leave out anything important.

1	Determine what you want to know and who can tell you.
2	Workshop the constructs and dimensions, then match to respondents.
3	Develop questions, matched to respondents, best question type, method to observe and develop data dictionary (determine strength of data type). Formulate your analysis plan.
4	Assemble survey, add introductions, definitions, references, add instructions and skip logic.
5	Develop plan for administration. Administer survey.
6	Carry out analysis and draft reports.

1	<p>Determine what you want to know.</p> <ol style="list-style-type: none"> 1. Develop 1 or 2 overarching questions. Keep them to a minimum and simple. <ol style="list-style-type: none"> a. An example of an overarching question could be: “What is the current emergency management (EM) capacity of urgent care providers in NYC overall?” b. And a follow up question could be: “What differences exist between urgent care provider types?” c. In developing your survey, always keep these questions in mind as your goal is to answer them through analysis of the answers you obtain from your survey. 2. Clearly state to yourself why you want to ask these questions. This gives your survey scope and prevents you from trying to measure the whole universe of possibilities in one <u>construct</u>. In this example, you may wish to answer the above questions <i>in order to determine gaps or places to focus on for future activities and funding</i>. 3. Determine what constructs you want to explore related to your overarching question. A <u>construct</u> is a something you want to measure but can’t measure directly like you could if it were an observable object or action (like correctly putting on PPE) or a reading on a device (like blood pressure). Emergency management capacity is a construct because many things make it up. <ol style="list-style-type: none"> a. You might see a few constructs in your questions above. One is <u>emergency management</u> (EM). 	<p>Checklist:</p> <ul style="list-style-type: none"> <input type="checkbox"/> 1 or 2 overarching questions <input type="checkbox"/> why these questions, to what end? <input type="checkbox"/> determine constructs <input type="checkbox"/> determine construct dimensions <input type="checkbox"/> determine persons knowledgeable about these dimensions <input type="checkbox"/> build in equity considerations <input type="checkbox"/> determine sampling frame <input type="checkbox"/> determine access to data that exists or you need respondents to report or calculate
---	---	---

A Guide to Evaluation for NYC HCC Programs

	<ul style="list-style-type: none"> b. Another is <u>capacity</u>. c. And finally, <u>sector providers</u>, as your audience, is a construct. d. Think of these constructs as a Venn diagram. Where they overlap is where your survey items and respondents come together. <ol style="list-style-type: none"> 4. Determine what dimensions make up these constructs. <ul style="list-style-type: none"> a. EM could be all the activities thought to go under this umbrella, like assessment, planning, training, exercising, communications and/or supplies. But it could be more. It takes workshopping (see next step) to figure this out. b. Capacity could be the definition of capacity. Going to the dictionary can then answer your question. But it could be “capability and capacity” which is how EM practitioners look at these terms. Capability is what you or a provider can do. Capacity is what you or a provider can do when pushed by bringing other resources into the frame. c. You may also need to explore EM and/or capability and capacity dimensions that come from regulatory bodies for your area of practice. 5. Determine what person or persons (respondent) is most likely to give you <u>usable</u> answers about these dimensions. 6. Bear in mind equity issues when developing constructs and respondents. It is advisable to have a diverse representation on the team developing the survey. 7. Keep in mind your sampling frame (or who or what you are questioning) serves your overarching questions. Sampling frames can be the <i>person</i> frame, the <i>organization</i> frame (or facility level). It could also be the <i>regional</i> frame. The lower/smaller the better most times, but this depends on scope, resources and your ability to reach those smaller frames. When you ask at the smallest, you can always aggregate to larger frames, but you can’t do the reverse when you ask at a large frame. 8. Lastly, think about what data already exists. For example, if you want percentages of growth or facility use, you will be asking your respondent to look this up and/or rely on reporting in their facility. If you want respondents to use/report or leverage this data, you must first know if respondents have access to this data and secondly, you will have to plan for more time for the respondent to answer. This will come up in your survey administration plan. 	
--	--	--

2	<p>Workshop the constructs and dimensions, then match to respondents.</p> <ol style="list-style-type: none"> 1. Gather colleagues who know your sector or domain for a discussion of constructs. They should be people who might respond (respondents) to your survey or help those who will. 	<p>Checklist:</p> <ul style="list-style-type: none"> <input type="checkbox"/> gather colleagues representative of respondents
---	---	---

A Guide to Evaluation for NYC HCC Programs

3	<p>You may include some colleagues from outside your sector or domain if you are going to discuss things that may be unfamiliar or new to the respondents, like “EM.”</p> <ol style="list-style-type: none"> 2. Ask them what they think of when they hear “EM,” “capacity,” “capacity,” and “urgent care providers.” By including your colleagues in this workshop, you close the gap between what you intend on a survey and what is understood by the respondents taking the survey. 3. Write down what they offer. Make a list under each construct of all ideas provided. 4. Ask if any of the ideas occur together or if they occur in a certain order. (You’ll use this for writing your questions later.) 5. Finally, normalize the lists. If similar ideas are offered that could be called one thing, see if the group agrees to simplifying them. For example, if under EM you have “PPE” and “masks,” ask if masks is the same thing as PPE. Break down PPE as much as you think is appropriate to the experience of those using PPE among your intended respondents. If similar or same ideas appear in lists under more than one construct, ask which is the best place for them or what the argument is for keeping them in both places. Using the PPE example again, you may find it listed under capability because the group you’ve gathered thinks of it as something staff are trained in. But it may also fall under EM if people think of it as materials or supplies. 6. Ask if a construct is something you need rated on a scale, reported exactly or if perception/opinion will tell you what you want to know. 7. (If you have time, validate the ideas that come to mind under each construct by giving the lists to other colleagues to see if they agree or would move or eliminate some of the ideas.) 	<ul style="list-style-type: none"> <input type="checkbox"/> workshop/brainstorm thoughts that come to mind about constructs <input type="checkbox"/> normalize the list <input type="checkbox"/> determine what kind of answer you want for each construct (scaled, exact number, perception/opinion) <input type="checkbox"/> validate the list
---	---	--

3

Develop questions, matched to respondents, best question type, method to observe and develop data dictionary (determine strength of data type). Formulate your analysis plan (pages 8-9 in this guide). Use this table. This will result in your survey questions and your analysis plan.					
Construct	Question for that construct (this is what goes into your survey)	Respondent match ¹	Question type ²	Define the variable.	How will you use the answer in relation to other answers?
Example: demographics of providers	Select your provider type: 1. Not for Profit 2. For Profit	This item is best answered by: ____	This is a categorical or nominal variable.	Provider type is the combination of two nominal variables describing for profit status and	We will use these variables to compare multiple groups when asking about differences in EM

A Guide to Evaluation for NYC HCC Programs

		3. Free Standing/Independent			structural dependency.	capability and capacity between and among them.
		4. Affiliated with a Health Care System				

4	<p>Assemble survey, add introductions, definitions, references, add instructions and skip logic.</p> <ol style="list-style-type: none"> 1. Use the items from the table above and place questions in a document (MS Word most likely). 2. Ensure items and item choices use consistent wording and format. Use consistent terminology and define terminology as needed. 3. Arrange the questions by topic and label each section with that topic as a header. Add text for context as needed for that heading. This will include definitions for terms respondents may not know or see often. 4. Ask yourself if the order makes sense to the respondents you plan to target to answer the survey or each section of the survey. When things don't make sense, rearrange. 5. Add any skip logic or skip directions needed. No responses to some answers will logically imply that further items on the same topic can be skipped by the respondents. 6. Test your survey with a small group of respondents similar to your target respondent group. Implement any changes they recommend for clarity. This will help you get data you can use when you administer with your target group. 7. Finally, add your introduction text. Tell your respondents why you are doing the survey, what you will do with the data, how their privacy is protected (confidentiality). And tell them when they can expect to see a report of the results. 	<p>Checklist:</p> <ul style="list-style-type: none"> <input type="checkbox"/> draft your survey <input type="checkbox"/> organize questions under headings (rearrange as needed) <input type="checkbox"/> add text for context, including definitions <input type="checkbox"/> add directions, skip logic or skip directions <input type="checkbox"/> test your survey and revised as needed <input type="checkbox"/> add the introduction text
---	--	--

5	<p>Develop plan for administration. Administer survey.</p> <ol style="list-style-type: none"> 1. See page 12 of our evaluation guide for how to increase your response rate and steps to follow to get your respondents on board. 2. Build in time for administration if the survey requires respondents to meet to answer it or if any items require them to review existing data or calculate data for survey responses. 3. We recommend that you dedicate time each week the survey is out to review questions from respondents and develop an FAQ so that you can prevent answer similar questions repeatedly. 4. Once you have the survey data back and the time for completion is over, make a note to the best of your ability about those who did NOT complete the survey. This will help in your limitations and interpretation of the results. And it may help you revise your 	<p>Checklist:</p> <ul style="list-style-type: none"> <input type="checkbox"/> develop your plan <input type="checkbox"/> administer the survey <input type="checkbox"/> monitor responses <input type="checkbox"/> develop FAQs and post them where all respondents can see them <input type="checkbox"/> note who has not completed the survey and write a brief paragraph describing them.
---	---	--

A Guide to Evaluation for NYC HCC Programs

	administration plan for the next time you do the survey or a similar one.	
--	---	--

6	Carry out analysis and draft reports. <ol style="list-style-type: none">1. Review and follow your analysis plan.2. Review page 7 of our deliverable quality guide on reporting writing.3. Interpret results.4. Include limitations (on methods and interpretation of results).5. Include next steps for re-assessment and a timeline.6. Include tasks revealed by the assessment and suggest coalition partners that can assist.	Checklist: <ul style="list-style-type: none"><input type="checkbox"/> follow your analysis plan<input type="checkbox"/> carry out data analysis<input type="checkbox"/> draft your report<input type="checkbox"/> interpret results bearing in mind why you did the survey (see item 1 above)<input type="checkbox"/> present your report and consider feedback where helpful to addressing your questions and why you did the survey (in item 1 above)
----------	--	--