INSTITUT NATIONAL DE SANTÉ PUBLIQUE DU QUÉBEC

Guide for the evaluation of a warning system for people vulnerable to heat and smog

Québec

Guide for the evaluation of a warning system for people vulnerable to heat and smog

Direction de santé environnementale et toxicologie

March 2013

INRS
Université d'avant-garde

Institut national de santé publique Québec

AUTHORS

Kaddour Mehiriz, Ph.D
Institut national de la recherche scientifique, Centre Eau Terre Environnement

Pierre Gosselin, MD, MPH
Institut national de santé publique du Québec
Institut national de la recherche scientifique, Centre Eau Terre Environnement
Centre de recherche et d'expertise en santé mondiale et environnement

LAYOUT

Julie Colas
Direction de santé environnementale et toxicologie
Institut national de santé publique du Québec

This document is available in PDF format on the Web site of the Institut national de santé publique du Québec at http://www.inspq.qc.ca.

LEGAL DEPOSIT – 1st QUARTER 2014
BIBLIOTHEQUE ET ARCHIVES NATIONALES DU QUEBEC
LIBRARY AND ARCHIVES CANADA
ISBN: 978-2-550-70134-7 (FRENCH PDF)
ISBN: 978-2-550-70229-0 (PDF)

ACKNOWLEDGMENTS

The authors wish to thank Health Canada's Climate Change and Health Office for their financial and scientific support in the preparation of this guide. The contributions of Abderrahmane Yagouti, Anna Yusa, Christina Daly, Frédéric Valcin, Jaclyn Paterson, Peter Berry from Health Canada and Sharon Jeffers from Environment Canada were appreciated. This work was also made possible by support from IRIACC-FACE and INSPQ.

SUMMARY

Weather warning systems are among the main policy tools used by public authorities to protect the population from vagaries in the weather. More particularly, weather warning systems are implemented so that the government can monitor weather conditions and issue warnings when extreme weather events, such as heat waves, cold waves, heavy rainfall, etc., threaten the population. As is the case in all public interventions, it is important to analyze the performance of these systems, in order to evaluate their contribution to the population's safety and well-being.

In this report, a guide for evaluating warning systems for people vulnerable to heat and smog is presented. It is intended for managers and evaluators of these systems. It provides them with an analytical framework and methods for evaluating their relevance, implementation and impacts.

The guide proposes that relevance analysis must address the need for warning systems for people vulnerable to heat and smog. More particularly, the emphasis should be on knowing whether the population in general, or some social groups in particular, require a warning system to be well informed about heat waves and smog episodes and about the behaviours to adopt to protect themselves from these vagaries.

It is also proposed in this guide that the implementation analysis focus on the capacity of the warning system to detect heat waves and/or smog episodes and to warn people who are vulnerable to them at the appropriate time. The quality of the warning messages, mainly the accuracy, utility and clarity of the embodied information, are also part of the implementation analysis.

For its part, the impact analysis investigates the effects[1] of heat and smog wave warning messages on the following:

1. Knowledge about the occurrence of heat waves and/or smog episodes, their consequences on health, as well as the most effective protective behaviours,
2. Attitudes towards the recommended behaviours during heat waves and/or smog episodes,
3. Perceptions of social norms relating to the recommended behaviours during heat waves and/or smog episodes,
4. Perceptions of the capacity to adopt the recommended behaviours during heat waves and/or smog episodes,
5. Intentions to adopt the recommended behaviours during heat waves and/or smog episodes,
6. Behaviours during phases of heat and/or smog, and finally,
7. The health status of vulnerable people exposed to heat waves and smog episodes.

[1] In the terminology used in this guide, effect, outcome and impact are synonymous.

The guide also proposes 28 indicators for evaluating heat waves and smog episodes warning systems. This list covers the main aspects of the evaluation, namely the analysis of the relevance, the analysis of the implementation, and the analysis of the impacts of heat waves and smog episodes warning systems.

The methods for evaluating heat waves and smog episodes warning systems are also included in this guide. Particular attention has been paid to the main techniques of impact evaluation, namely the experimental methods, the difference-in-difference estimator, propensity score matching, panel data analysis, time series analysis, and the instrumental variables. The strengths and weaknesses of each of these techniques are presented.

Finally, the evaluation of warning systems is considered, in this guide, to be a complex operation that requires advanced knowledge of evaluation methods as well as the capacity to adapt them to the particular context of each system. Considerable funds are also necessary to finance the different evaluation activities, such as the development of the evaluation protocol, the collection and analysis of data, and the dissemination of the evaluation results. The support of public health authorities is also needed to properly carry out a comprehensive evaluation of warning systems as suggested in this guide.

TABLE OF CONTENTS

INTRODUCTION

The average global temperature has risen over the last century and it is expected that this warming will be accompanied by more frequent and increasingly intense heat waves and smog episodes (Buset *et al*., 2008; Das *et al*., 2012). As shown by the heat waves in Europe in the early 2000s, exposure to extreme heat and smog constitutes a real threat to the health of the population. Some social groups, particularly the elderly or those individuals suffering from respiratory or cardiovascular diseases, are more vulnerable to heat and smog.

There are currently numerous definitions for heat waves that take into account characteristics of the climate and the vulnerabilities of countries or regions in the world. Environment Canada defines a heat wave as a period of three or more consecutive days when the maximum temperature is 32 degrees Celsius or higher. In hot regions of the world, a heat wave is defined as involving temperature thresholds higher than the Canadian threshold. This threshold is established at 40 degrees Celsius in Australia and 45 degrees Celsius in India (Das, 2012).

In numerous situations, heat waves are accompanied by smog, generally defined as a high concentration of atmospheric pollutants (Government of Canada, 2012; Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs du Québec, 2007). Atmospheric pollutants constitute a threat to the health of the population, since they are the cause of many respiratory and cardiovascular diseases (World Health Organization (WHO), 2002). Furthermore, their effects are aggravated during heat waves (Buset *et al*., 2008).

Weather warning systems (WWS) are a fundamental part of adaptation to weather hazards including heat waves and smog episodes. The purpose of a WWS is to provide relevant information on weather hazards so that individuals and communities can plan for and react at the right time and in the appropriate way to reduce negative impacts on the population's health and property and on the environment (Bacher, 2006; United Nations, 1997).

WWS are intended to support decision makers to fulfill the four following functions (Bacher, 2006; Gosselin *et al*., 2012; United Nations, 1997 and 2006):

1. Analyze the risks and effects of exposure to threatening climatic conditions;
2. Observe, detect, monitor, analyze and anticipate threatening weather conditions;
3. Disseminate warning messages in a timely, reliable and understandable way to the authorities and people in danger;
4. Plan locally the response to emergencies, and prepare and train populations for an appropriate response to the alerts in order to reduce the impacts of the threatening conditions.

Similar to many other developed countries, Canada has had a national weather warning system for a long time. Environment Canada is the national agency responsible for weather forecasts and for issuing weather watches and warnings when extreme climatic events are likely to affect the safety or property of Canadians. Since the early 2000s, Canada has also seen the creation of several WWS at the provincial level, as is the case for Québec, which has implemented a warning system for people vulnerable to heat and atmospheric pollutants (SUPREME project) (Toutant *et al*. 2011). Some Canadian cities such as Montréal and Toronto also have their own WWS.

The guide presents an analytical framework and a set of methods for evaluating warning systems for heat waves and summer smog episodes in Canada. The analytical framework is used to select the most relevant evaluation questions as well as to develop the appropriate indicators that will be used. The methodology consists of determining, based on the evaluation questions, the design that will be used, the sources of data, as well as the techniques of data collection and analysis.

1 APPROACH FOR GUIDE DEVELOPMENT

The approach used for development of the WWS evaluation guide for heat wave and summer smog episodes consisted of three main steps. In the first step, a literature review was conducted in order to enhance and update understanding of the risks associated with heat waves and smog, and to identify the most vulnerable people, as well as the means for reducing these risks. In addition, relevant theoretical models of human behaviour were reviewed, in order to better understand the effects of WWS on the behaviour of people vulnerable to heat and smog. The theories and methods of program evaluation used for evaluating WWS were also examined. From this review, best practices in the field of WWS evaluation were identified, specifically the evaluation questions most commonly addressed, as well as methods and outcome measurement indicators used.

In the second step, the authors participated in the committees of two WWS research and development projects. They were the Vigilance project (partnership between the Institut national de santé publique du Québec (INSPQ), Environment Canada, and the Ministère de la Sécurité publique du Québec) and a project on weather warnings for people vulnerable to heat and smog (partnership between the INSPQ, the Agence de santé et services sociaux de la Montérégie, the Centre de santé et de services sociaux Pierre-Boucher, and the urban agglomeration of Longueuil). Expert input from representatives from Health Canada was also solicited in order to clearly understand the needs of the guide’s users. Regular participation in the work of weather warning project committees as well as the consultations that took place with Health Canada increased the authors understanding of the objectives and operation of WWS in addition to the issues involved in evaluating these systems.

The third step involved validation of the guide by four experts from Health Canada and Environment Canada.

2 DEFINITION AND PRINCIPLES OF PROGRAM EVALUATION

This section presents the definitions and the key principles of program evaluation that informed the development of this guide.

Program evaluation is the systematic use of research approaches and methods in order to arrive at a credible judgement about the value of public interventions (Rossi *et al*., 1999). As an integral part of the public management process, program evaluation contributes to decision-making by analyzing the relevance, implementation and impact of public action.

The aim of relevance evaluation is to analyze the program's *raison d'être*. The starting point is the principle that public funds are rare resources that must be used carefully to meet the population's real needs. Implementation evaluation emphasizes the program's implementation and operation. It generally focuses on the analysis of its resources, processes and outputs (e.g. warnings). The main issue at this level is to determine whether the program's goods and services have been offered as initially planned for the intended people and organizations, and at the lowest cost. The aim of impact evaluation is to measure the intended and unintended outcomes (e.g. reduction of heat-related mortality) of the program.

Evaluation is expected to provide reliable information on the relevance, effectiveness and efficiency of public programs. In fact, by analyzing the reasons that led to the creation of the program, a judgement can be made about its relevance. The effectiveness of a program is based on a comparison of its outcomes and objectives. The results of impact evaluation therefore would be used to assess to what extend the program has met its objectives. Finally, the comparison between the program outcomes and resources allows a judgement to be made about its efficiency. The question of knowing if the outcomes are worth the cost incurred is the fundamental issue at this level of analysis.

3 ANALYTICAL FRAMEWORK OF HEAT WAVES AND SMOG EPISODES WARNING SYSTEMS

The purpose of the present analytical framework is to delimit the scope and evaluation criteria for a WWS system for people vulnerable to heat and/or smog. The framework addresses three major questions, namely:

1. Does the WWS meet a well-identified social need requiring the government's intervention?
2. Does the WWS function properly?
3. Does the WWS effectively contribute to protecting people vulnerable to heat waves and smog episodes?

In program evaluation, these questions correspond to the analysis of relevance, the analysis of implementation, and the analysis of impacts of WWS, respectively.

3.1 ANALYSIS OF THE RELEVANCE OF HEAT WAVE AND SMOG EPISODE WARNING SYSTEMS

The goal of WWS is to provide the necessary information on weather hazards as well as on the recommended means of protecting the health and property of vulnerable people. Needs analysis emphasizes these two particular aspects of WWS for people vulnerable to heat and/or smog. More particularly, a WWS would be needed if the evidence shows the presence of a heat wave and/or smog episodes exposure risks and that the general population or specific social groups are not well informed about the occurrence of these events, their effects and the protective actions to be taken. The new system thus fills a need for information that could not have been met otherwise. This aspect is very important in evaluating the relevance of WWS, because in numerous countries, there are several agencies that monitor and disseminate weather information by a multitude of means of communication, such as radio, television, newspapers, etc. This is why most studies show that the population is generally well informed about extreme weather hazards (Environment Canada, 2001 and 2012). Consequently, it becomes important to know what features of the new WWS allow it to give added value to the existing warning systems. Does the new system stand out, due to its greater capacity to reach vulnerable social groups, to disseminate the information, to fit the information to the needs of the users, etc.?

3.2 ANALYSIS OF THE IMPLEMENTATION OF HEAT AND SMOG WARNING SYSTEMS

Implementation analysis emphasizes the primary functions of a WWS, namely providing reliable and relevant information on a timely basis to those individuals vulnerable to heat waves and smog episodes. The analysis therefore should address the following aspects of implementation: the reliability of weather forecasts, the relevance of eligibility criteria for receiving warning messages, the relevance of warning thresholds, the capacity to reach the target population and finally, the quality of the warning messages. These are presented below.

3.2.1 The reliability of weather forecasts

WWS provide weather forecasts intended for people and organizations in order to help them take weather into consideration in planning their activities. To be useful in decision-making, these forecasts need to be reliable (CRUE, 2008). This is why comparing forecasts to observed weather conditions is generally used to evaluate the quality of WWS forecasts (Kovat and Ebi, 2008; CRUE, 2008).

3.2.2 The relevance of eligibility criteria for heat and smog warnings systems

A number of weather warnings are intended for people who are more likely to suffer from heat waves and smog because of pre-existing medical conditions. These include:

1. People 65 years of age and older;
2. People with heart and/or lung problems;
3. People suffering from renal insufficiency;
4. People suffering from diabetes or a neurological disease;
5. People suffering from a mental health disorder.

Other social determinants, like poverty and low level of education, may be of importance as they constitute a barrier to preventive measures such as air conditioning during heat waves (Reid *et al*. 2009). Certain neighbourhood characteristics, such as the presence of green spaces or important urban heat islands, might alleviate or worsen exposure to hazards.

In some cases, the organizations responsible for WWS do not possess the necessary scientific knowledge on the impacts of heat and smog on population health, nor the data and the capability to generate this knowledge. If this is the case, warnings may be ineffective in targeting the most vulnerable groups, which ultimately leads to poor performance of the WWS.

3.2.3 The relevance of heat and smog warning thresholds

Warning messages are generally disseminated when weather forecasts show that temperature and/or the concentration of air pollutants will be higher than some predefined levels which are considered to be harmful to the population health. Implementation analysis therefore should focus on the relevance of weather warning thresholds, mainly the question of whether identified thresholds for alert systems are established on the basis of a rigorous analysis of the effect of heat and smog on the population's health.

The population's behaviour must also be considered in determining heat and smog warning thresholds. In this regard, public health authorities responsible for WWS are faced with the following dilemma. First of all, establishing warning thresholds at a lower temperature level or atmospheric pollutant concentration has the advantage of warning, and hopefully protecting the health of, the most vulnerable people in society. However, the lowering of warning thresholds inevitably leads to an increase in their number. In this case, users may not seriously take them into account, because they associate them with weather conditions that, in their opinion, are not a threat to their health. An increase in the number of messages may then produce the opposite effect, namely a reduction in their impact on the behaviour of the population. Determining the optimal warning threshold is therefore a challenge for public health authorities responsible for WWS.

3.2.4 The capacity to reach vulnerable individuals to heat and smog

Warning messages are generally intended for people living in a specific territory and presenting certain characteristics that, according to the available knowledge, make them vulnerable to heat and smog. Once the eligibility criteria have been established, the organization must have the required knowledge, technology and resources to have their warning messages reach people vulnerable to heat and smog. In many situations, this capacity is lacking or not sufficient at best (Abrahamson and Raine, 2009; Maibach *et al*., 2008; Polivka *et al*., 2012). For instance, organizations would simply not have updated lists of people vulnerable to heat and smog. The WWS may also not possess the technological infrastructure as well as the human resources necessary for it to send warning messages to vulnerable individuals. These problems are likely encountered in the case of warnings that are not intended for the general population, but for specific social groups such as the elderly suffering from chronic diseases. In this case, these individuals must be identified and their contact information be available so that the appropriate mechanisms are established to reach them, such as telephone calls, home visits or the distribution of brochures (Health Canada, 2012).

The capacity to reach individuals who are vulnerable to heat and smog remains a concern because the effectiveness of the WWS in reducing health risks depends on it (Bassil and Cole, 2010). Implementation evaluation must therefore focus on the percentage of the people vulnerable to heat and smog who have in fact (1) received warning messages and (2) understood the contents of these messages. A WWS cannot have an impact on the behaviour of individuals if they have not received and understood the alerts.

3.2.5 The quality of heat and smog warning messages

Experts in warning systems agree that for meteorological alerts to be effective, they must include information on the occurrence of a threatening weather event as well as advice to help people adopt appropriate behaviours to protect their health and property (United Nations, 2006). This information must also be communicated sufficiently in advance of the weather event so that people have enough time to take appropriate action.

The effectiveness of the message does not depend solely on its content and timing, but also on its understandability for the intended recipients (Kalkstein and Sheridan, 2007). This means that the message must be formulated clearly and adapted to the different audiences of the WWS. The evaluation of the quality of heat and smog warning messages should therefore take into account all the above elements.

3.3 ANALYSIS OF THE IMPACTS OF HEAT AND SMOG WARNING SYSTEMS

The goal of this section is to identify the expected impacts of weather warning messages on people vulnerable to heat and smog. To do this, the theoretical foundations of the identification strategy are presented. The theoretical model will then be used to establish the list of expected outcomes of issuing heat and smog alerts as well as the confounding factors that, in addition to warning messages, have an influence on the achievement of these outcomes.

3.3.1 The theoretical framework of impact analysis

WWS are based on the assumption that individuals do not have sufficient information about the occurrence of threatening weather events, the effects of these events, and the behaviours required to optimally protect their health and property. It is also assumed that individuals will adopt the appropriate behaviour following their exposure to warning messages and information issued by weather services (Kalkstein and Sheridan, 2007; United Nations, 2006) and that this change in behaviour in turn will reduce the harmful effects of the weather hazards on the population's health and well-being (Das, 2012).

However, studies show that exposure to awareness campaigns in general, and to weather alert messages in particular, does not always lead to a change in behaviour (Kalkstein and Sheridan, 2007; Sheridan, 2007; Snyder, 2007). Several factors may weaken the effect of the message on behaviour, such as the poor quality of the message itself, the high cost of adopting the recommended behaviour, an unreceptive attitude towards such messages, etc.

Warning messages do not affect behaviour directly, but only indirectly by means of their impacts on factors that shape the individuals' behaviour, such as their perception of their vulnerability to heat waves and smog episodes, their attitude towards the recommended behaviour, the social pressure to adopt a given behaviour, etc. On this subject, several theoretical models have been developed in order to understand the factors that explain behaviour. Among the principal public health models, the health belief model, the theory of reasoned action, and finally, the theory of planned behaviour (Bélanger and Godin, 2003) are particularly relevant. Recent work has focused on the synthesis of these models in order to develop new ones that suitably account for explanatory factors of behaviour (Fishbein, 2008; Yzer, 2012). These models are also used in the design and evaluation of public health programs (Yzer, 2012).

The guide proposes using the integrative model of behavioural prediction (Fishbein, 2008; Yzer, 2012) to identify the expected effects of WWS on people vulnerable to heat and smog. This model is based on the idea that the behaviour of individuals is determined in large part by their intentions to carry out the behaviour in question, such as reducing outdoor activities when it is very hot. Fishbein (2008) considers that besides intentions, one must also take into account the environmental constraints in which the behaviour will be adopted, as well as the actual capacity (and not the perceived capacity) to adopt it. Indeed, even if individuals intend to adopt a protective behaviour and think that they have the means to do it, the actual constraints of the environment, such as the absence of a cooling system that can be used during heat waves, may preclude them from acting according to their intentions. As individuals may not have all the information about their environment, nor the time and capacities necessary to process this information thoroughly, perceptions may underestimate or overestimate the actual possibilities and constraints in the environment.

Intention to act depends on the attitudes, perceived norms and perceived self-efficacy. Attitudes refer to a general feeling that is either favourable or unfavourable to the adoption of a specific behaviour. Perceived norms refer to the effects of perceived pressure from the social environment on the individuals' behaviour. The model assumes that individuals take into consideration in their behaviour the viewpoint of the people who are most important in their eyes, such as family members and co-workers. Finally, self-efficacy refers to what extent the individuals feel capable of adopting the behaviour. Since the idea is that even if the individuals have a favourable attitude about a given behaviour and feel supported by their social network, they may reject this behaviour if they consider that they are unable to adopt it.

Attitudes, perceived norms and self-efficacy are in turn the product of a set of beliefs that find their roots in more deeply established factors such as socioeconomic context, gender and the culture of the individuals.

The integrative model of behavioural prediction has the advantage of being a parsimonious model that proposes a small number of variables that explain a good part of behaviour. This model also allows development of the causal chain by which a public health intervention would impact the population's health and well-being, particularly in the case of interventions using the provision of information as a means of behavioural change, as is the case with weather warning systems. Moreover, this model is widely used in designing and evaluating programs (Yzer, 2012; Mattern *et al*., 2011; Bleakley *et al*., 2011; Hill *et al*., 2007).

3.3.2 The expected effects of heat and smog warnings

The integrative model of behavioural prediction was used to develop a logic model of heat wave and smog warning systems. The logic model is presented in Figure 1 below. The green rectangle designates the outputs of the WWS. The blue rectangles refer to the expected effects of the warning messages on people. For their part, the grey rectangles refer to confounding factors that may be correlated with the variables used to measure the effect of the WWS on the individuals.

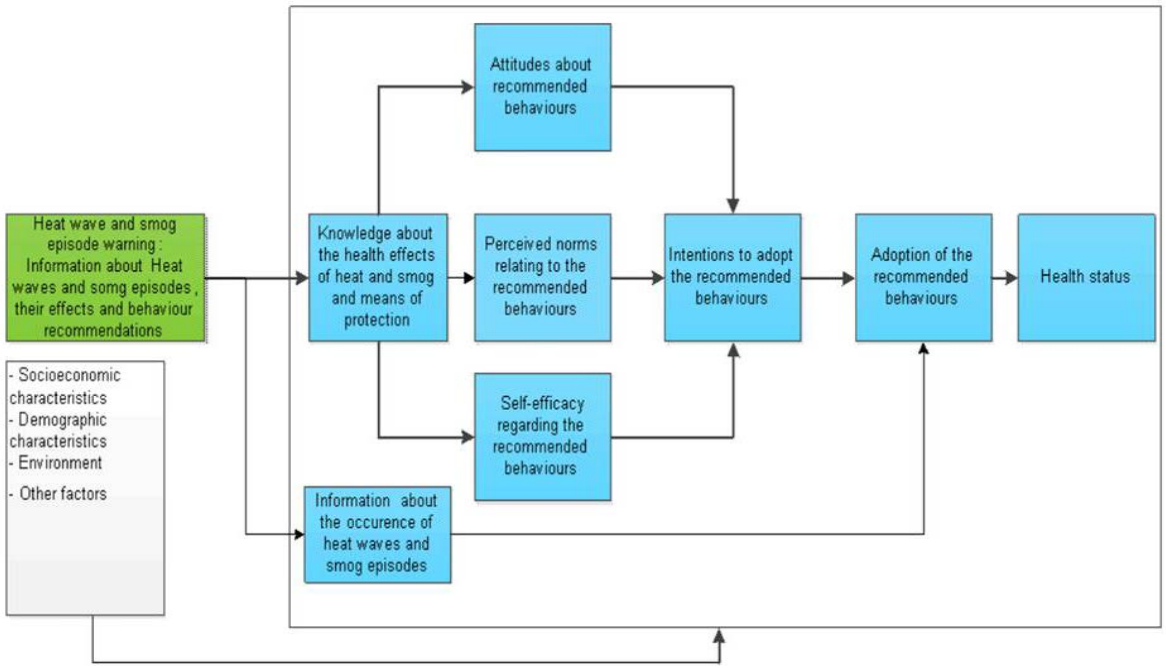

Figure 1 The logic model of heat and smog warning system

According to the logic model, WWS for people vulnerable to heat and smog would thus have the seven categories of effects described below.

3.3.2.1 Effects on knowledge

The objective of WWS is to disseminate information about threatening weather conditions, the risks that ensue from these events, and to provide advice to protect people against them. It is assumed that warning messages let people to be well informed on the occurrence and health consequences of heat waves and smog episodes and on the effective protective behaviours (e.g. drinking more water) that should be adopted.

Studies show that most people are generally well informed about the occurrence of heat waves and, to a lesser extent, about smog episodes. However, it seems that a significant proportion of the population does not change its behaviour during these episodes, mainly due to the underestimation of the harmful effects of heat waves and smog episodes on health as well as a lack of knowledge about preventive measures and doubts about their effectiveness (Bassil and Cole, 2010; Wolf *et al*., 2011).

This is why it is useful to know the extent to which the warning messages resulted in the people being better informed about (1) the arrival of heat waves and smog episodes, (2) health risks from exposure to these events, and (3) the required adaptation strategies.

3.3.2.2 Effects on attitudes

The improvement in knowledge about heat waves and smog episodes following exposure to warning messages is expected to lead to favourable attitudes and ultimately, the adoption of the recommended behaviours. The idea is that warning messages change beliefs about the harmful effects of heat waves and smog as well as the effectiveness of the behaviours recommended for reducing them.

3.3.2.3 Effects on perceived norms

Perceived norms may have an impact on the intention to adopt protective measures against heat waves and smog episodes. Wolf *et al*. (2010) show that support networks have an influence on the perceptions and behaviours of the elderly during heat waves. The expectations and standards conveyed by medical personnel can also have an impact on the behaviour of people vulnerable to heat and smog.

3.3.2.4 Effects on the self-efficacy

Warning messages generally have a set of recommended actions to reduce the harmful effects of heat and smog, such as drinking a sufficient amount water (heat) or reducing activities that require intense physical effort (heat and/or smog). These messages may also indicate to people the resources made available to them by the community during heat waves or smog episodes, such as access to air-conditioned premises or to public pools. The knowledge that results from exposure to weather warning messages should therefore have the effect of improving people's perceptions about their ability to adopt behaviours that appropriately protect them from heat waves and smog.

3.3.2.5 Effects on intentions

It is expected from the integrative model of behavioural prediction that the favourable effects of warning messages on attitudes, perceived norms and self-efficacy would increase the intention of vulnerable people to adopt appropriate behaviours during heat waves and smog. The evaluator should therefore ask to what extent exposure to the warning messages changed the intention of the individuals to adopt the desired behaviours. The importance of documenting this aspect is due to the fact that intentions are the main determinants of behaviour.

3.3.2.6 Effects on behaviour

People who are well informed about the occurrence of heat waves and/or smog episodes and have the firm intention to adopt recommended behaviour would be more likely to undertake measures to protect themselves from related adverse effects.

The procedure generally used to assess the effects of warnings on people's behaviour consists of establishing a list of behaviours recommended in the warning messages and of subsequently asking the respondents to what extent they did or did not adopt these behaviours (Kalkstein et Sheridan, 2007). The list of behaviours to be monitored must therefore be based on an attentive examination of the warning messages for people vulnerable to heat and smog.

3.3.2.7 Effects on health

Due to its intended effects on people's behaviour, a well-performing warning system is likely to contribute to the reduction in the number of cases of morbidity and mortality caused by heat waves and smog episodes.

Effects on morbidity

Studies on this subject generally use the number of hospitalizations and emergency room visits as measurement indicators of the effect of heat and smog on the health status of the population (Brunekreef and Holgate, 2002; Knowlton, 2009; Kovat and Ebi, 2006; Mastrangelo, 2007). However, the relationship between heat and smog episodes and their impact on population health is not well established, and the results of several studies are not consistent (Knowlton, 2009). This relationship could be contingent on several other factors such as the population's socioeconomic characteristics, the prevalent health status, the actual exposure to the risks, the availability of air conditioning, access to green spaces, etc.

Effects on mortality

Several studies have analyzed the effect of heat and/or smog on mortality. The studies generally focus on people 65 years of age and older who are considered as being the most vulnerable to heat waves and smog episodes (Kovat and Ebi, 2006). The few published studies seem to show that WWS tend to reduce the number of deaths as well as morbidities caused by heat (Bassil and Cole, 2010; Chau and Woo, 2009; Das, 2012).

Before presenting the confounding factors, it is worth noting that the effects of warning messages depend in large part on the nature of these messages. Some messages would intend for example to change people's perception of the adverse effects of heat waves and smog episodes (e.g., general preventive message at the start of the season), or their perception of self-efficacy in implementing preventive measures. The resulting change in perceptions leads to changes in attitudes, intention and behaviour at a later stage. However, these messages would not have an impact on the perception of social norms. Particular attention must therefore be paid to the content of the message as well as to its objectives when the time comes to establish the list of indicators that will be used to monitor WWS outcomes.

3.3.3 The confounding factors

The model suggests that in addition to exposure to warning messages, other factors also have an impact on the behaviour of individuals as well as on their health status.

3.3.3.1 People's characteristics

Socioeconomic and demographic characteristics, cultural factors, living environment, etc., of people vulnerable to heat and smog would have an impact on their attitudes, perceived norms, self-efficacy, intentions, behaviours and health status. Semenza *et al*. (2008) suggest that in the USA, individuals with lower income and educational attainment and non-whites are more likely to be aware of the degradation of air quality and extreme heat and to protect themselves from these hazards.

3.3.3.2 The environment

The constraints of the social and natural environment are also among the determinants of behaviour and health status. As an example, a number of warning messages recommend to people that they cool their homes by opening windows in the evening or by turning on air conditioning systems. The elderly with low incomes may decide not to follow any of these recommendations, because on the one hand, they do not have the financial means for air conditioning, and on the other, they fear that leaving windows open during the night increases the risks of intruders into their homes.

4 METHODS FOR EVALUATING HEAT AND SMOG WARNING SYSTEMS

In the first part of this document, a framework for evaluating WWS for people vulnerable to heat and smog was proposed. According to this framework, the evaluation should be based on three main components, namely analysis of the relevance, the implementation, and the effects of the WWS. This part of the guide is devoted to the presentation of a list of indicators as well as the methodologies that can be used to inform data collection and ultimately to conduct the evaluation.

4.1 INDICATORS FOR EVALUATING HEAT AND SMOG WARNING SYSTEMS

Based on the analytical framework put forward in this guide, Table 1 proposes a series of generic indicators that would be used to evaluate heat waves and smog episodes warning systems.

Table 1 Indicators for the evaluation of WWS to protect people vulnerable to heat and smog

Aspects of the evaluation		Indicators
Relevance analysis	Risk of heat wave and smog episode exposure in the area of concern	1. Probability of occurrence of heat waves and smog episodes 2. Intensity of heat waves and smog episodes (level of exposure) 3. Duration of heat waves and smog episodes (days)
	Health effects of heat waves and smog episodes in the area of concern	4. Presence of scientific evidence showing heat waves and smog episodes represent a threat to the population's health
	Population awareness of heat waves and smog episodes	5. The extent to which the population is informed on the occurrence and the adverse effects of heat waves and smog episodes and on the protective behaviour
Implementation analysis	Quality of the weather forecasts	6. Percentage of false positive/negative heat and smog warnings
	Relevance of eligibility conditions for receiving heat and smog warnings	7. Warning messages are intended to individuals who, according to scientific evidence available, are most vulnerable to heat waves and/or smog episodes
	Relevance of warning thresholds	8. Appropriateness of scientific evidence on which the thresholds are based 9. WWS users' perception of the relevance of warning thresholds

Table 1 Indicators for the evaluation of WWS to protect people vulnerable to heat and smog (cont'd)

Aspects of the evaluation		Indicators
Implementation analysis (cont'd)	Capacity to reach people vulnerable to heat waves and smog episodes	10. Percentage of eligible individuals who received heat and smog warning messages
	Awareness of the heat and smog warning messages	11. Percentage of recipients of heat and smog warning messages who listened to the warning message
	Quality of the heat and smog warning messages	12. Understandability of the messages 13. Relevance of the information contained in the messages 14. Appropriate time between the warning and the occurrence of the hazard 15. Satisfaction with the quality of the warning messages
Impact analysis	Effects of warning messages on knowledge	16. Awareness of actual heat waves and smog episodes 17. Knowledge about the health effects of heat waves and smog episodes 18. Knowledge about the behaviours to be adopted for protection against heat waves and smog episodes
	Effects of warning messages on attitudes	19. Perception of the effectiveness of the recommended behaviours 20. Perception of the inconveniences associated with taking the recommended behaviours
	Effects of warning messages on perceived norms	21. Respondents' perception of the expectations of their social network (parents, friends etc.) regarding the adoption of the recommended behaviours
	Effects of warning messages on self-efficacy	22. Perception of the capacity to adopt the recommended behaviours 23. Perception of the barriers preventing the adoption of the recommended behaviours
	Effects of warning messages on intentions	24. Intention to adopt the recommended behaviours in future heat waves and smog episodes
	Effects of warning messages on behaviour	25. Adoption of the recommended behaviours
	Effects of warning messages on the population's health	26. Number of visits to Emergency rooms 27. Number of hospitalizations for relevant diseases 28. Number of deaths for relevant death causes

One should note that this list is not exhaustive and that indicators adapted to the characteristics of each WWS must be developed. As an example, the recommended behaviours may be different, depending on the characteristics of the people to which they are addressed. The managers of one WWS may retain behaviours that the managers of another system consider less relevant for protecting health. Also, the evaluation needs not be based on all the proposed indicators. If the focus is on the impact of warning messages on population health, only information on the health of the participants in the study is required, such as the number of visits to Emergency rooms and the number of deaths.

The exact formulation of each indicator must also be adapted to the evaluation methodology used. Thus, in an evaluation using a randomized design to assess the effects of warnings on the behaviour (section 5.4.1), it is sufficient to ask the respondents whether they adopted the recommended behaviours or not. A frequency scale, going from ‘rarely’ to ‘very often’ for example, can be used for this purpose. Afterwards, the average score for experimental and control groups is calculated, and the difference between the two groups is used to estimate the effects of warnings on behaviour. Under some circumstances, the evaluator can rely on the self-reported effects by the people who received warning messages. In this case, the percentage of people who report that the warning messages had an impact on their behaviour could be used as an outcome indicator. Finally, several qualitative indicators are proposed that are more suitable for the analysis of relevance and implementation of WWS.

The indicators listed in Table 1 should be used to collect the needed information for evaluating the relevance, implementation and impacts of heat wave and smog warning systems. It should be noted in this regard that the evaluator may use existing tools in developing his/her own data collection instruments. A questionnaire is found in the appendix of the Kalkstein and Sheridan (2007) study that has been used by the authors to measure the effects of heat warning systems. Environment Canada has also developed questionnaires that are used for evaluating its weather services in general and warning services in particular. These questionnaires are, however, not designed on the basis of the integrative model of behavioural prediction. Consequently, they are not intended to measure several expected effects of WWS, such as the effects on attitudes, perceived norms, intentions, etc. Readers should also consult the guides designed to support researchers using this model, or the theory of planned behaviour, especially the guides of Azjen (2002) and Gagné and Godin (1999), both of which are available online.

4.2 METHODS FOR ANALYZING THE RELEVANCE OF WWS

Relevance analysis has the objective of establishing whether there is a need for a heat and smog warning system, in a given location. So, it will be important to identify if:

1. The area has been subject in the past, or could be subject in the future to heat waves and smog;
2. Heat wave and smog episodes have negative effects on the population’s health;
3. The population is not well informed about heat waves and smog occurrence, their effects on health, and/or the behaviours to adopt in such situations.

To document these aspects, the evaluator should first address the reasons that led public health authorities to implement the heat waves and smog episodes warning system. This can be achieved through an analysis of the project’s framework documents in which the system’s managers present the problem, the options considered, as well as the solution chosen. Interviews with the system’s managers will also be used to collect information about the rationale for implementing a WWS. The evaluator could also do a more thorough analysis by trying to determine whether managers' need assessment in fact corresponds to the population's needs. Various tools could be used for this, such as surveys of WWS users, the consultation of studies on the impacts of heat waves and smog on population health in the area deserved by the WWS, etc.

4.3 METHODS FOR ANALYZING THE IMPLEMENTATION OF WWS

The implementation analysis focuses on the reliability of the weather forecasts, the relevance of weather warning eligibility criteria (for the warnings targeting vulnerable groups), the relevance of weather warning thresholds, the ability to reach vulnerable people, and finally, the quality of the heat and smog warning messages.

Environment Canada's archives contain information for comparing real conditions with weather forecasts, and from this, allow reliable information to be obtained about the reliability of the forecasts of heat waves and smog episodes.

The effectiveness of weather warnings depends on their ability to identify and enroll people most vulnerable to heat and/or smog. Interviews with WWS designers will provide information about the relevance of the eligibility criteria used for the selection and enrollment of the recipients of the warning messages. This information can be completed by consulting the scientific literature or by interviews with caregivers for people vulnerable to heat and smog, experts in this field, etc. Here, the issue involves knowing whether the social groups that are actually more vulnerable to heat waves and smog episodes are eligible for weather warnings.

As for reaching the target audience and the quality of the messages, most of the evaluations currently available conduct surveys on people vulnerable to heat and smog in order to determine whether they have received the weather warnings. Questions (that vary with the dissemination support for these messages) are also asked to collect information about the quality of the warning such as the accuracy, the relevancy, the understandability and the timeliness of the messages (Ekos, 2011; Huppé *et al*., 2013; Semenza *et al*., 2008).

4.4 METHODS FOR ANALYZING THE OUTCOMES OF WWS

Establishing a causal link between people's exposure to heat wave and smog warning messages and changes in their knowledge, behaviours, health status, etc., is the fundamental issue in evaluating the effects of WWS. From an experimental standpoint, "A cause is viewed as a manipulation or a treatment that brings about a change in the variables of interest, compared to some baseline, called the control" (Dehedjia and Wahba, 2002). A WWS has the effect of reducing mortality if the number of deaths in the group of people who have received warning messages is less than the number of deaths that would have occurred if the members of this same group have not been exposed to these messages. The problem that arises here is that it is impossible to have outcome data at a given point in time on the situation with and the situation without warnings messages. The unavailability of data results from the fact that a given person can either receive or not receive a warning message; that person cannot be in both situations at the same time. The evaluator is therefore confronted with the problem of finding a valid counterfactual that will be used to estimate the effect of weather warnings on the recipients.

Several methods have been developed to overcome this problem, with the most important ones being random assignment, propensity score matching, the difference-in-difference estimator, panel data analysis, time series analysis, and finally, instrumental variables. As is shown below, these methods that are not mutually exclusive use outcomes data from two or more groups, from the same group at different times or a combination of the two techniques to assess the effect of public interventions. Moreover, evaluators often use them in a complementary manner to improve the rigour of their evaluations, mainly to test the robustness of their findings to the methods used.

4.4.1 Random assignment

Random assignment consists of randomly creating two groups of people vulnerable to heat and smog. The members of the first group, called the experimental group, will receive heat and smog warning messages, while the members of the control group will not receive them. Belonging to either of the groups is by chance, so the two groups are in principle equivalent, and hence, the control group is a valid basis for depicting the situation that would have prevailed without the heat wave and smog warning system.

When outcomes data are compiled, relatively simple statistical analyses (such as mean difference tests between the experimental group and the control group, and ANOVA) are used to measure the effect of exposure to heat wave and smog warnings. Based on these calculations, it can then be concluded that a WWS reduces the number of deaths caused by heat and smog when the proportion of people in the experimental group who had died during these episodes is less than that in the control group.

However, it should be noted that random assignment is not always feasible, due to practical, political or ethical considerations. Thus, this model is difficult to use in evaluating weather warnings that are disseminated by the media and that consequently are available to the general public. In this type of intervention, the evaluator has no control over the people who will receive and those who will not receive the messages. Public health authorities may also be reluctant to randomly assign vulnerable people to experimental and control groups because individuals who did not receive warning messages could be considered less protected from the harmful effects of heat wave and smog than the recipients of these messages.

Impact evaluations are confronted to the problem of selection bias whenever random assignment methods have not been used. Selection bias refers to "… processes and events not under the researcher's control that lead some members of the target population to be more likely than others to participate in the program under evaluation […] Such pre-existing differences, when related to outcome variables, are known as selection bias" (Rossi *et al.*, 1999: 241). Selection bias is therefore present when the experimental group is not equivalent to the control group. It is a serious threat to the internal validity of the evaluation of the WWS outcomes. This is the case, for example, when the experimental group consists of people who are less vulnerable to heat than the control group. Comparison of the two groups would result in this case in erroneous conclusions due to these pre-existing differences. More particularly, the evaluator may note that the proportion of deaths in the experimental group is smaller than that in the control group, and then attribute this difference to the weather warnings, while in reality, this difference between the two groups existed even before the WWS was implemented. The impact evaluation methods presented in the following sections are generally used to attenuate selection biases.

4.4.2 Propensity score matching

The aim of matching methods is to correct for selection biases that are due to observable differences between the participants and non-participants in a WWS (Dehejia and Wahba, 2002). Observable differences refer to the characteristics of the members of the two groups that the evaluator is able to measure in the context of a WWS evaluation, such as health status and the perception of the heat and smog risk. Matching consists of creating a control group that is similar to the experimental group[2], for a series of variables that explain people's exposure to the heat wave and smog warning messages. A control group could be created on the basis of the criteria that were used to select the recipients of warning messages, such as the age of the recipients, the presence of cardiovascular and respiratory diseases, income, etc. When the eligibility criteria are not well defined, the control group could also be created by taking into account factors that determine vulnerability to heat and smog. The aim of this exercise is to create a control group whose members have the same likelihood of being selected to receive the warning messages than experimental group members.

The difficulty in finding for each recipient of warning messages an individual who did not receive the messages but has the same characteristics is a major challenge in this method. It becomes practically impossible when there are several factors that are used to recruit the recipients of heat and smog warnings (Becker and Ichino, 2002).

Propensity score matching is intended to solve this problem (Rosenbaum and Rubin, 1983). The propensity score refers to "[…] the conditional probability of assignment to a particular treatment given a vector of observed covariates" (Rosenbaum and Rubin, 1983: 41). This technique is being increasingly used in program evaluation (Fu *et al*., 2006; Nguyen *et al*., 2006; Trujillo, Portillo and Vernon, 2005). It consists of matching the participants and non-participants in the WWS on the basis of their propensity score for receiving warning messages instead of on their characteristics.

The propensity score is a more effective matching method for creating a control group. It matches people based on a single variable, namely the propensity score, instead of the vector of variables that was used to calculate this score.

There are three steps in the propensity score matching process. The first step involves selecting a sample of recipients and non-recipients of weather alerts and using logistic regression techniques or a probit to calculate the probability of each of the members in the

[2] We are aware that the use of the terms experimental group and control group is not totally appropriate in the case of non-experimental methods. To simplify the text, we will nevertheless use the term experimental group to designate the participants in the study who will receive the warning messages, and the term control group to designate the people participating in the study who will not receive the warning messages.

sample receiving these warnings. For this, the evaluator must choose a list of relevant variables, such as heat and smog risk factors or eligibility criteria, to predict participation in the WWS.

Then, ideally each recipient of the WWS must be matched with a non-recipient who has the same propensity score. The goal is to create a control group whose members have the same propensity scores as the members of the experimental group. However, the propensity score is a continuous variable, and consequently, it is practically impossible to find two observations that have the same propensity score (Becker and Ichino, 2002). To overcome this problem, several methods have been developed, with the most commonly used being the following: *Nearest-Neighbor matching method, stratification method, radius matching method,* and *Kernel matching method* (Becker and Ichino, 2002).

The final step involves comparing the experimental group and the control group on the basis of the WWS outcomes measures, such as the proportion of people who are aware of heat waves, the proportion of people who were hospitalized during this episode, etc.

The propensity score matching method is based on the hypothesis that propensity score matching allows to select a control group that is equivalent to the experimental group. The idea is that after having controlled the observable factors that explain WWS participation, any noted difference between the two groups is due either to exposure to the warning messages, or to chance. In most cases, this hypothesis is difficult to defend due to the existence of non-observable differences between the two groups, even after having taken carefully into consideration the recipient selection criteria. In fact, in most telephone weather alerts, the heat wave and smog warnings are aimed at people who, on the one hand, meet the eligibility criteria, and on the other, agree to receive these messages. At that time, several factors that motivate people to participate in telephone warnings are not accessible by the evaluator, and when these factors are correlated with the outcomes measurement indicators, they constitute a threat to the internal validity of the evaluation of the impact of WWS. For example, this is the case when, among the people eligible for a heat and smog warning system, those who consider themselves very vulnerable to these phenomena are more likely to sign up for the WWS than the other members of the group. If this difference in perception of the heat and smog threats is not taken into consideration in calculating the propensity score, the effects of warning messages on behaviour can be overestimated because the people who receive these messages naturally tend to adopt the recommended behaviours.

4.4.3 Difference-in-difference estimator

The existence of data about the situation before issuing heat wave and/or smog warning messages allows the difference-in-difference estimator to be used to reduce the selection biases that are not observable (Abadie, 2005; Heckman *et al*., 1997; Khandker *et al*., 2010). As is the case with matching, the difference-in-difference estimator requires that an experimental group and a control group of people vulnerable to heat and smog be formed. The first group consists of people who receive warning messages, and the second group, of people who do not receive them.

To calculate the effect of the WWS, one first has to calculate within each group the difference between the situation before and after issuing the heat wave and/or smog warning messages. The measurements may involve, for example, the degree of compliance to a list of behaviours that are recommended by health authorities during heat waves and smog. The result obtained is used in a second step to compare the two groups, particularly to determine whether the people who received the warning messages are more likely to adopt the recommended behaviours than those individuals who did not receive any messages. Any noted difference is then attributed to the warning messages.

The difference-in-difference estimator is based on the hypothesis that the differences between the two groups that are not due to the WWS remain constant over time, whether these differences are observable or not (Abadie, 2005; Heckman *et al*., 1997). Using the previous example, this model produces unbiased results when the pre-existing behavioural differences between the two groups are constant over time. However, when these differences change naturally over time, they may bias the estimation of the effects.

Propensity score matching techniques and the difference-in-difference estimator can be combined to deal with the selection biases caused by observable factors that vary over time, as well as selection biases that are due to non-observable factors but that are constant over time (Heckman *et al*., 1997; Abadie, 2005). Consequently, this method is more rigorous than propensity score matching or the difference-in-difference estimator taken separately (Abadie, 2005; Heckman *et al*., 1997). However, it does not control for the time sensitive non-observable differences between experimental and control groups.

4.4.4 Panel data analysis

Panel data analysis is a generalization of the difference-in-difference estimator (Khandker, 2010). More precisely, this method can be used when a person does not systematically receive warning messages during heat waves and smog. In this case, a longitudinal database can be created on the basis of surveys of vulnerable people during heat waves and smog, from the data on the health status of these individuals, etc. This database then allows a comparison of the behaviour of the person with her/himself, but also with other people.

Panel data analysis has the advantage of controlling for the selection biases caused by the factors that vary over time, but that are observable, meaning the variables included in the regression model, as well as the non-observable factors but that are constant over time, such as gender, ethnic origin, etc. As a result, they are among the most rigorous impact evaluation designs. However, this technique does not allow control of the non-observable differences between the two groups, which change over time.

A model for panel data analysis was used by Das *et al*. (2012) to estimate the effect of a program designed to prevent mortality due to heat waves in the state of Odisha in India during the 1998-2010 period. The author compared over 12 years the number of deaths caused by heat in the 17 districts that benefited from the program to the number in the 13 districts that did not benefit, while taking into account the socioeconomic characteristics and heat waves in each district. The results of the study indicate that the program reduces the number of deaths caused by heat.

4.4.5 Time series analysis

Time series are generally used to estimate the effect of heat on mortality (Doyon *et al*., 2006; Fateh *et al*., 2012; Fouillet *et al*., 2007). They involve following, over a long period, the relationship between heat waves and the number of deaths recorded in a given territory. These techniques are also used to evaluate the effects of weather warnings on the reduction of the number of deaths caused by heat waves.

To illustrate the use of time series analysis, we will take the case of a country that implemented a WWS in 2010 and that wants to estimate the effect of this system on the number of deaths caused by heat waves. The procedure consists of estimating the excess mortality caused by this hazard by using the data of the twenty or thirty years preceding 2010. This estimate is used as counterfactual, meaning as a measurement of the number of excess mortalities that would have been recorded in the absence of WWS during the years 2010 to 2013. Subsequently, the excess mortality from heat waves is estimated for the years 2010-2013, a period during which the WWS was put into service. Comparison of the two estimations is used to calculate the effect of the WWS on the reduction in deaths caused by heat waves (Fouillet *et al*., 2008; Kalkstein *et al*., 2011).

This method has two essential shortcomings. The first involves the validity of the estimations of the excess mortality caused by heat. The technique used at this level consists of comparing the number of deaths during heat waves with the number of deaths during normal periods. Since heat waves and periods of normal temperature do not occur at the same time, the evaluators are faced with the problem of finding a valid counterfactual for estimating the effect of heat waves on the number of deaths. The second problem involves the validity of using data from the past as a baseline for comparing the excess mortality of future periods. Vulnerability to heat waves and smog tends to decrease with the evolution in technology, the improvement in the standard of living, as well as the development of knowledge on the health effects of these hazards and the means of reducing them. The difference between the two periods may not be due uniquely to the implementation of the WWS, but also to the development of the adaptation capacity of the individuals and of the society as a whole, such as improved emergency or nursing care services.

4.4.6 Instrumental variables

The objective of instrumental variable techniques is to minimize the biases caused by non-observable factors. The principle underlying this technique consists of using, as instruments, variables that meet the two following conditions. First, the instrumental variables must have a good capacity for predicting participation in weather warning systems. Second, these variables must not have an impact on the outcomes variables, but only through their influence on participation in WWS. In the absence of the WWS, the instrumental variables are not deemed to be correlated with the outcomes variables. As an example, while age predicts well those individuals who will receive warning messages (condition 1), it cannot be used as an instrument. The reason is because age is also among the factors that explain the mortality caused by heat waves (violation of condition 2). The difficulty of finding variables that fulfill these two conditions is the main limitation of this technique (Trujillo *et al*., 2005).

4.5 PROGRAM EVALUATION IS THE ART OF THE POSSIBLE

In the preceding sections, several methods for evaluating the effects of WWS on reducing the adverse effects of heat waves and smog episodes on the population health have been proposed. However, it should be noted that the capacity to use these methods depends on the particular context of each evaluation. It is therefore possible to use random assignment to evaluate weather warning pilot projects. Decision-makers can randomly choose sites in which WWS will be tested, or within a specific site, the people who will benefit from these warning messages. The choice of evaluation methods is however very limited in the case of WWS that are already implemented. The creation of experimental and control groups is possible when the WWS covers only some of the people or territories that are vulnerable to heat and smog. This is the case, for example, when a health organization decides on its own to establish a heat and smog warning system. Under these conditions, it is possible to compare the behaviour of people who have received warning messages with that of a control group made up of the population served by a health organization that has not established a warning system. However, since assignment to experimental and control group is not random, simple comparison between groups is not sufficient to reliably measure the effects of WWS. Other methods must also be considered to control for the selection bias, such as measurements before and after the implementation of the WWS, and the creation of the control group based on the WWS recipient selection criteria. For some outcomes, archives are available that make it possible to obtain measurements of the situation before and after the implementation of WWS, even if this information was not collected specifically for the evaluation. This is the case for archives relating to the number of hospitalizations and deaths, and heat waves which are used by several longitudinal studies on the relationship between heat waves and the population's health.

Under certain circumstances, none of the above-proposed methods are feasible. In this case, one can resort to a set of less rigorous evaluation methods, but while being aware of their weaknesses. A number of evaluations are therefore limited to conducting surveys of the people who received warning messages, in the days following a heat wave or smog episode (Sheridan, 2007; Kalkstein and Sheridan, 2007). Among others questions, the respondents are asked to report to what extent their behaviour has been influenced by weather warning messages. These methods are based implicitly on the assumption that the individuals are able to establish a causal link between their exposure to the messages and their behaviour. More precisely, it is assumed that the people are able 1) to determine correctly the behaviour that they would have adopted without the warning messages, 2) to compare it to the actual behaviour, and 3) to faithfully report the results of this comparison. However, the individuals' cognitive limitations can lead to less precise estimations of the effects of weather warnings. People who are concerned about the survival of the WWS may also be tempted to consciously overestimate the positive effects and underestimate the negative effects of warning messages (strategic behaviour).

Faced with the impossibility of creating a control group, the evaluator can also compare the situation before and after the implementation of a WWS. The noted difference can be attributed in this case to the heat waves and smog episodes warnings. But one must be aware that this method is based on the assumption that in the absence of the WWS, the individuals' behaviour remains the same. Behaviour changes over time. More particularly, aging can have the effect of increased awareness of heat and smog vulnerability, and as a result, a greater likelihood of adopting the behaviours that protect against their negative health effects.

Finally, it is worth noting that health authorities need to have a well-developed evaluation capacity to successfully implement the aforementioned methods. This includes hiring personnel with high-level skills in program evaluation methods, notably the qualitative and quantitative techniques of data collection and analysis, and the communication of evaluation results. Significant financial resources are also needed to cover the costs of evaluation activities. Moreover, a strong commitment of senior management to evaluation is a valuable asset that enables the organization to efficiently use its evaluation capacity to produce reliable information on the pertinence, implementation and outcomes of its WWS. When evaluation capacity is insufficient, health authorities can outsource the evaluation of its WWS to organizations that possess the required skills and resources such as universities and private firms specialized in program evaluation or policy analysis.

CONCLUSION

Global warming is accompanied by more intense and more frequent heat waves and smog than before. Studies indicate that heat waves and smog are the cause of excess mortality, mainly among the elderly with chronic diseases and from deprived neighbourhoods. Mortality caused by heat and deterioration in air quality has over time become a public health problem requiring the government's intervention. In this respect, a number of public health authorities, both nationally and locally, have implemented WWS as elements in their climate change adaptation strategy. Faced with the increased use of WWS, it becomes important to have analytical frameworks and methods for evaluating the contribution of these tools to the protection of the population against the harmful effects of heat and smog.

In this document, we have presented a guide for evaluating warning systems to protect people vulnerable to heat and smog. More precisely, we have proposed an analytical framework that would serve as a basis in determining the issues as well as the indicators for evaluating WWS. We have also presented the main evaluation methods that can be used for this. As was mentioned above, the indicators and methods used must be adapted to the specific context of each evaluation. Depending on the WWS, some indicators may prove to be less relevant, while some methods cannot be used. An evaluator therefore has the difficult task of choosing and implementing the best methods that remain available, given the opportunities and the constraints of the evaluation context.

REFERENCES

Abadie, A. (2005) Semiparametric difference-in-difference estimators. *The Review of Economic Studies*. 72: 1-19.

Abrahamson, V. and Raine, R. (2009) Health and social care responses to the Department of Health Heatwave plan. *Journal of Public Health*. 31 (4): 478-489.

Ajzen, I. (1985) From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.). Action-control: From cognition to behavior. Heidelberg, Germany, Springer: 11-39.

Ajzen, I. (2002) Constructing a TpB Questionnaire: Conceptual and Methodological Considerations. Available from http://www.unibielefeld.de/ikg/zick/ajzen%20construction%20a%20tpb%20questionnaire.pdf. Page consulted on April 25, 2013.

Ajzen, I. (2011) Job satisfaction, effort, and performance: A Reasoned action perspective. *Contemporary Economics*. 25 (4): 32-43.

Armitage, C. and Conner, M. (2001) Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology*. 40: 471-499.

Basher, R. (2006) Global early warning systems for natural hazards: systematic and people-centered. *Philosophical Transactions of the Royal Society*. 364: 2167-2182.

Bassil, K.L and Cole, C.D. (2010) Effectiveness of public health interventions in reducing morbidity and mortality during heat episodes: a structured review. *International Journal of Environmental Research and Public Health*. 7: 991-1001.

Basu, R., Dominici, F. and Samet, J.M. (2005) Temperature and mortality among the elderly in the United States: A comparison of epidemiologic methods. *Epidemiology*. 16 (1): 58-66.

Becker, S. and Ichino, A. (2002) Estimation of average treatment effects based on propensity scores. *The Stata Journal.* 2 (4): 358-377.

Bélanger, D. *et al*. (2006) *Vagues de chaleur au Québec méridional: adaptations actuelles et suggestions d'adaptations futures*. INSPQ.

Bélanger, D. and Godin, G. (2003) La psychologie sociale au service de la santé publique et de l'environnement. In Michel Guérin *et al*.: *Environnement et santé publique: Fondements et pratiques*. Edisem: 277-290.

Bleakley, A., Hennessy, M., Fishbein, M. and Jordan, A. (2011) Using the Integrative Model to Explain How Exposure to Sexual Media Content Influences Adolescent Sexual Behavior. *Health Education Behavior*. 38 (5): 530-540.

Brunekreef, B. and Holgate, S. (2002) Air pollution and health. Lancet. 360: 1233–42

Buset, C. *et al*. (2008) Air Quality, Climate Change and Health. In Séguin, J. *Human Health in a Changing Climate: A Canadian Assessment of Vulnerabilities and Adaptive Capacity*. Ottawa, Health Canada: 129-192.

Bustinza, R., Toutant, S., Gosselin, P. and Bélanger, D. (2011) S*ystème intégré des données de vigie et surveillance des aléas hydrométéorologiques et géologiques (SIDVS-AHG)*. Évaluation de l'utilisation du volet chaleur au cours de l'été 2010. INSPQ.

Chau, P.H., Chan, K.C. and Woo, J. (2009) Hot weather warning might help to reduce elderly mortality in Hong Kong. *International Journal of Biometeorology*. 53: 461-468.

Chebana, F., Martel, B., Gosselin, P., Giroux, J. and Ouarda, T. (2012) A general and flexible methodology to define thresholds for heat health watch and warning systems, applied to the province of Québec (Canada). *International Journal of Biometeorology.* DOI 10.1007/s00484-012-0590-2.

CRUE (2008) Effectiveness, efficiency and early warning system for flash-floods. Research report n1-5. Available from http://www.crue-eranet.net/images/EWASE_factsheet.pdf. Page consulted April 24, 2013.

Das, S. and Smith, S. (2012) Awareness as an adaptation strategy for reducing mortality from heat waves: Evidence from a disaster risk management program in India. *Climate Change Economics.* 3 (2):1-29.

Dehejia, R. and Wahba, S. (2002) Propensity score-matching methods for non-experimental causal studies. *The Review of Economics and Statistics.* 84 (1): 151-161.

Doyon, B., Bélanger, D. and Gosselin, P. (2006) *Effets du climat sur la mortalité au Québec méridional de 1981 à 1999 et simulations pour des scénarios climatiques futurs.* Québec, Institut national de santé publique.

Ebi, K., Teisberg, T., Kalkstein, L., Robinson, L. and Weiherheat, R. (2004) Watch/Warning Systems Save Lives. Estimated costs and benefits for Philadelphia 1995-98. *American Meteorological Society*: 1067-1073.

Environnement Canada (2001) *Opinion publique sur les avertissements météorologiques*. Rapport final.

Environnement Canada (2011) *Sondage de 2011 sur la qualité du service rendu par les Services météorologiques et environnementaux.* Rapport final.

Environnement Canada (2012) *Sondage national de 2012 sur les services météorologiques et environnementaux.* Rapport final.

Fassiatou, O., Bélanger, D. and Gosselin, P. (2010) *Proposition d'indicateurs aux fins de vigie et de surveillance des troubles de la santé liés à la chaleur.* INSPQ.

Fishbein, M. (2008) A Reasoned Action Approach to Health Promotion. *Medical Decision Making*, Nov-Dec: 834-844.

Fishbein, M. and Ajzen, I. (1975) Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, MA, Addison-Wesley.

Fouillet *et al*. (2008) Has the impact of heat waves on mortality changed in France since the European heat wave of summer 2003? A study of the 2006 heat wave. *International Journal of Epidemiology.* 37: 309-317.

Fu, A., William, H. and Gordon, G. (2006) Propensity score and difference-in-difference methods: A study of second-generation antidepressant use in patients with bipolar disorder. *Health Services Outcomes Research Methods.* 7 (1-2): 23-38.

Gagné, C. and Godin, G. (1999) L*es théories sociales cognitives: Guide pour la mesure des variables et le développement de questionnaire.* Available from http://www.godin.fsi.ulaval.ca/Fichiers/Rapp/Guide_mesure_variables.pdf. Page consulted April 25, 2013.

Giguère, M. and Gosselin, P. (2006) Événements climatiques extrêmes et santé. Examen des initiatives actuelles d'adaptation aux changements climatiques au Québec. INSPQ.

Gosselin, P. *et al*. (2012) *Vigilance météorologique santé*. Description du projet conjoint Environnement Canada-INSPQ.

Government of Canada (2013) *Canadian Smog Science Assessment – Highlights and Key Messages*. Library and Archives Canada cataloguing in publication. ISBN 978-1-100-19064-8

Government of Canada (2009) *Policy on Evaluation.* Available from http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=15024. Page consulted April 25, 2013.

Hajat, S. and Kosatsky, T. (2010) Heat-related mortality: a review and exploration of heterogeneity. *Journal of Epidemiology and Community Health.* 64: 753-760.

Health Canada (2012) *Heat Alert and Response Systems to Protect Health Best Practices Guidebook.*

Heckman, J., Ichimura, H. and Todd, P.E. (1997) Matching as an econometric evaluation estimator: Evidence from evaluating a Job Training Program. *The Review of Economic Studies*. 64: 605-654.

Hill, C., Abraham, C. and Wright, D.B. (2007) Can theory-based messages in combination with cognitive prompts promote exercise in classroom settings? *Social Science & Medicine.* 65: 1049-1058.

Huppé, V., Nguon, S. and Courteau, M. (2013) *Service automatisé d'alertes téléphoniques de la Cote air santé : étude de l'observance des recommandations de santé transmises chez un groupe de patients vulnérables à la qualité de l'air. Institut National de Santé Publique.* Available from http://www.inspq.qc.ca/pdf/publications/1646_CoteAirSante_ObservRecommSante.pdf. Page consulted September 30, 2013.

Kalkstein, A.J. and Sheridan, S.C. (2007) The social impacts of the heat-health watch/warning system in Phoenix, Arizona: assessing the perceived risk and response of the public. *International Journal of Biometeorology*. 52: 43-55.

Kalkstein, L.S., Scott, G., Mills, D.M. and Samenow, J. (2011) An evaluation of the progress in reducing heat-related human mortality in major U.S. cities. *Natural Hazard*. 56: 113-129.

Khandker, S., Koolwal, G. and Samad, H. (2010) Handbook on impact evaluation. Quantitative methods and practices. Washington, *The World Bank*.

Knowlton *et al*. (2009) The 2006 California heat wave: Impacts on hospitalizations and emergency department visits. *Environmental Health Perspectives*. 17 (1): 61-67.

Kovat, S.R. and Ebi, K.L. (2006) Heatwaves and public health in Europe. *European Journal of Public Health*. 16 (6): 592-599.

Kysely J. and Kříž, B. (2008) Decreased impacts of the 2003 heat waves on mortality in the Czech Republic: an improved response? *International Journal of Biometeorology*. 52: 733-745.

Maibach, E.W. *et al*. (2008) Climate change and local public health in the United States: Preparedness, programs and perceptions of local public health department directors. *Climate Change & Public Health*. 3 (7): 1-8.

Mastrangelo, G., Fedeli, U., Visentin, C., Milan, G., Fadda, E. and Spolaore, P. (2007) Pattern and determinants of hospitalization during heat waves: an ecologic study. *BMC Public Health.* 7: 1-8.

Matterne, U., Diepgen, T.L. and Weisshaar, E. (2012) A longitudinal application of three health behaviour models in the context of skin protection behaviour in individuals with occupational skin disease. *Psychology and Health*. 26 (9): 1188-1207.

Ministère de la Santé de des Services Sociaux (2013) *Projet pilote de la Cote air santé au Québec*. Available from http://www.coteairsante.qc.ca/. Page consulted April 24, 2013.

Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs (2007) L'origine du smog. *Info Smog*. Available from http://www.mddep.gouv.qc.ca/air/info-smog/fiche-form.pdf. Page consulted April 24, 2013.

Nguyen, A., Taylor, J. and Bradley, S. (2006) The Estimate effect of catholic schooling on educational outcomes using propensity score matching. *Bulletin of Economic Research.* 58 (4): 285-307.

Polivka, B.J., Chaudry, R.V. and Crawford, J.M. (2012) Public health nurses' knowledge and attitudes regarding climate change. *Environmental Health Perspectives.* 120 (3): 321-325.

Reid, C. *et al*. (2009) Mapping community determinants of heat vulnerability. *Environmental Health Perspectives.* 117 (11): 1730-1736.

Rosenbaum, P. and Rubin, D. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika.* 70 (1): 41-55.

Rossi, P., Freeman, H. and Lipsey, M. (1999) Evaluation – A Systematic Approach. Thousand Oaks, *Sage Publication.*

Semenza *et al*. (2008) Public perception and behavior change in relationship to hot weather and air pollution. *Environmental Research*. 107: 401-411.

Sheridan S.C. (2007) A survey of public perception and response to heat warnings across four North American cities: an evaluation of municipal effectiveness. *International Journal of Biometeorology*. 52: 3-15.

Smoyer-Tomic, K. and Rainham, D. (2001) Beating the heat: Development and evaluation of a Canadian hot weather health-response plan. *Environmental Health Perspectives.* 109 (12): 1241-1248.

Snyder, L.B. (2007) Health communication campaigns and their impact on Behavior. *Journal of Nutrition Education and Behavior* . 39:S32-S40

Toutant, S., Gosselin, P., Bélanger, D. and Rivest, S. (2011) An open source web application for the surveillance and prevention of the impacts on public health of extreme meteorological events: The SUPREME system. *International Journal of Health Geographics.* 10 (39): 1-11.

Trujillo, A., Portillo, J. and Vernon, J. (2005) The Impact of subsidized health insurance for the poor: Evaluating the Colombian experience using propensity score matching. *International Journal of Health Care Finance and Economics*. 5: 211-239.

United Nations (1997) *Report on early warning for hydrometeorological hazards including drought.*

United Nations (2006) *Developing early warning system: A check list.* The third international conference on early warning system.

United Nations (2012) Early warning systems: A State of the art analysis and future directions. ISBN: 978-92-807-3263-4. Available from http://www.crue-eranet.net/images/EWASE_factsheet.pdf. Page consulted April 24, 2013.

Wilson, L; Giles, A. (2013) A new index for the verification of accuracy and timeliness of weather warnings. *Meteorological applications.* 20: 206–216.

Wolff, K., Nordin, K., Brun, W., Berglund, G. and Kvale, G. (2011) Affective and cognitive attitudes, uncertainty avoidance and intention to obtain genetic testing: An extension of the Theory of Planned Behaviour. *Psychology and Health*. 26 (9): 1143-1155.

World Health Organization (2002) *Quantifying Selected Major Risks to Health*. Available from http://www.who.int/whr/2002/en/whr02_ch4.pdf. Page consulted April 24, 2013.

World Meteorological Organization (2008) *Capacity assessment of national meteorological and hydrological services in support of disaster risk reduction.*

Yzer, M. (2012) The Integrative Model of Behavioral Prediction as a tool for designing health messages. In Hyunyi Cho: *Health communication message design. Theory and practice*. Sage Publication: 21-40.

Publication N^{o}: 1794

Institut national
de santé publique
Québec