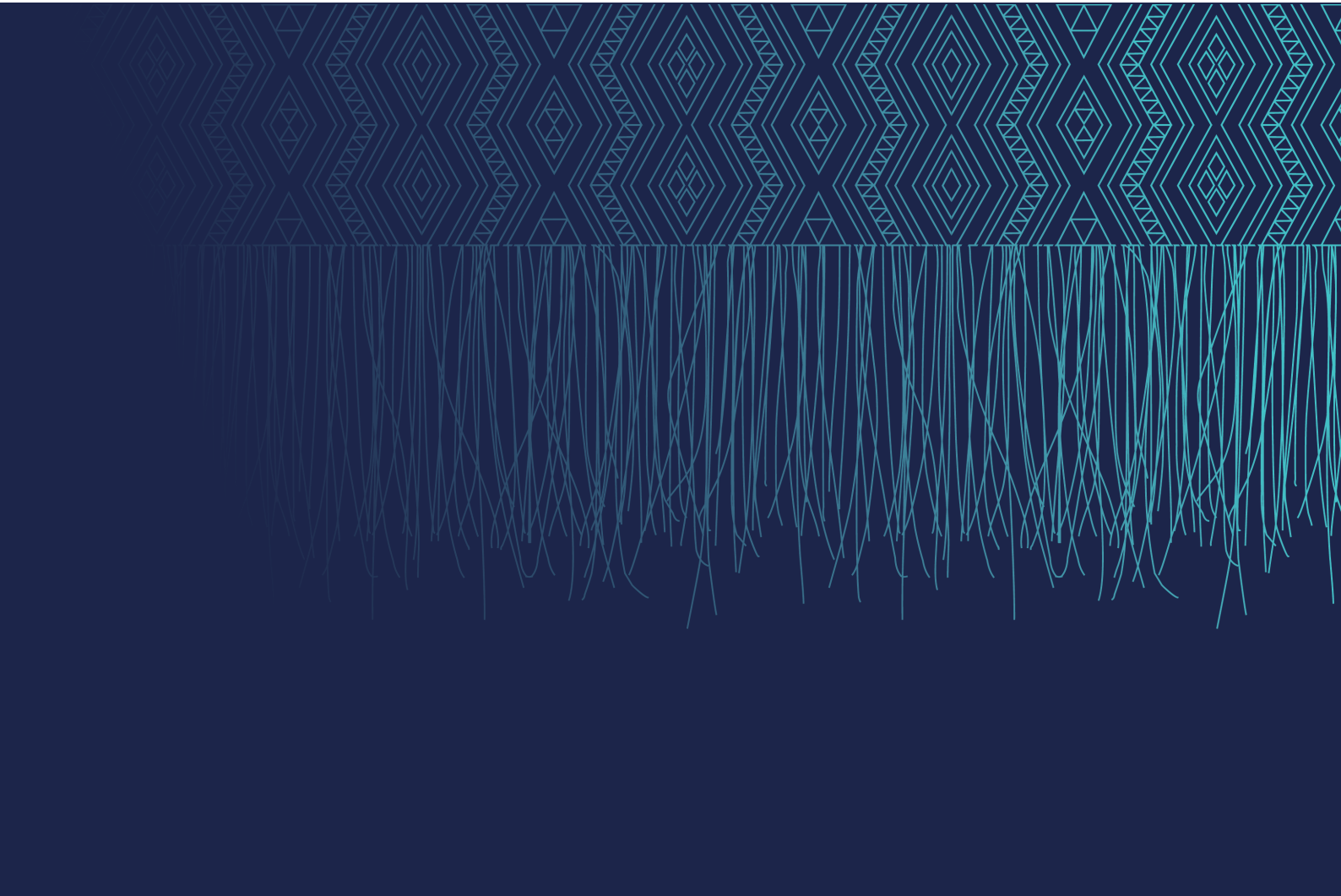


# Investigating Clusters of Non-Communicable Disease

Guidelines for Public Health Officers



Citation: Te Whatu Ora – Health New Zealand. 2023. *Investigating Clusters of Non-Communicable Diseases: Guidelines for Public Health Officers*. Wellington: Te Whatu Ora – Health New Zealand.

This document has been revised from the version published in May 2015 by the Ministry of Health, PO Box 5013, Wellington, New Zealand.

Published in November 2023 by Te Whatu Ora – Health New Zealand  
PO Box 793, Wellington 6140, New Zealand

ISBN 978-1-99-106774-6 (online)

**Te Whatu Ora**  
Health New Zealand

This document is available at [tewhatauora.govt.nz](https://www.tewhatauora.govt.nz)



This work is licensed under the Creative Commons Attribution 4.0 International licence. In essence, you are free to: share ie, copy and redistribute the material in any medium or format; adapt ie, remix, transform and build upon the material. You must give appropriate credit, provide a link to the licence and indicate if changes were made.

# Preface

The authors have revised *Investigating Clusters of Non-communicable Disease: Guidelines for public health services*, which was published by the Ministry of Health in 1997 and has been available online since 1998. The revised edition takes account of recent literature and organisational and legislative changes.

**Deborah Read** Public health physician

**Barry Borman** Epidemiologist

# Acknowledgements

The following acknowledgements apply to the 1997 edition.

We wish to acknowledge the assistance of Associate Professor Neil Pearce, Wellington School of Medicine, and Dr John Harris, Director, California Birth Defects Monitoring Programme, for peer review.

Comments were also provided by Dr Martin Tobias and Henry Dowler, Ministry of Health.

We also wish to thank Dr Gillian Durham, and medical officers of health who gave us feedback on a previous version of the guidelines produced by the Public Health Commission.

# Contents

<b>Preface</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>Executive summary</b>	<b>v</b>
<b>Background</b>	<b>6</b>
What is a cluster?	7
Examples of non-communicable disease clusters	7
Clusters in New Zealand	10
Surveillance for clusters	11
Occupational clusters	12
Cluster investigation and causation	12
<b>The cluster investigation process</b>	<b>14</b>
Stage 1: Preliminary evaluation of a report of an alleged cluster	16
Stage 2: Verification of index case and exposure reports	21
Stage 3: Full case ascertainment	26
Stage 4: Surveillance or epidemiological study	36
<b>Legislation</b>	<b>37</b>
<b>Ethics</b>	<b>38</b>
<b>Risk communication</b>	<b>40</b>
Perception of risk	41
Communicating with the public	41
Risk comparison	44
Relationships with the media	45
Conclusion	46
<b>References</b>	<b>47</b>
<b>Glossary</b>	<b>53</b>
<b>Appendix: Cluster report form</b>	<b>54</b>

# Executive summary

The investigation of alleged clusters of non-communicable disease, often prompted by public concern, can be a complex and resource-intensive activity that requires thorough planning and careful implementation.

These guidelines provide a systematic approach to such investigations. Public health officers can follow this approach to carry out an organised and coordinated response to reports of alleged clusters.

Investigation involves four distinct stages. Each successive stage involves collecting more specific data and requires a stronger verification of those data. However, a public health officer may choose to combine stages, depending on local judgement, experience and the available resources. A decision as to whether to proceed with further investigation is made at the end of each stage.

Stage 1 is the response when an alleged cluster is initially reported to the National Public Health Service. The procedure involves recording the initial report, developing a case definition and following up with the informant.

Investigating a cluster suspected from monitoring or vital statistics begins at Stage 2, when the index case(s) and exposure(s) are verified. Specific tasks include deciding who should carry out the verification, literature review, and identification and review of the appropriate records.

Stage 3 identifies the confirmed cases in the time period and geographical area of interest and determines if there is a statistically significant cluster. A case-finding team is formed, the case definition is revised, cases are identified and data collected, and the observed number of cases is compared with the number of cases expected in the time period and geographic area being investigated.

Stage 4 provides the option of continuing the investigation by either surveillance or an epidemiological study.

Dealing with the concerns of the public and media is fundamental to investigating clusters. The public often feel threatened about the occurrence of alleged clusters and demand action and information from the health professionals involved in an investigation. Skilled risk communication, understanding of risk perception and effective handling of inquiries from the public and the media are crucial to the success of a cluster investigation.

# Background

Although many more organisations have published guidelines for investigating non-communicable disease clusters since the Ministry of Health – Manatū Hauora published its original guidelines in 1997, they are largely based on earlier guidelines, including those of the United States Centers for Disease Control and Prevention (CDC 1990). The CDC has published an addendum and specific cancer cluster guidelines and toolkit to be used with its original guidelines (Kingsley et al 2007; Abrams et al 2013; National Public Health Information Coalition and CDC 2013). Examples of recent guidelines include European Surveillance of Congenital Anomalies (EUROCAT 2007), Alberta Health Services (2011), Queensland Health (2012) and the National Health and Medical Research Council (2012) in Australia.

The science and response to investigating clusters have changed little since 1997. The main change has been that more disease incidence and exposure (eg, biomonitoring) data and analytical methods and software, including geographic information systems (GIS), are now available.

The need to follow up reports of alleged clusters of non-communicable disease has strengthened with increasing public awareness and concern about certain environmental exposures. This task often has to be done in the glare of publicity and under urgent and stressful circumstances.

Although Rothman (1990) maintains cluster investigations have little scientific value, others consider they should be viewed on more than their scientific merit alone. Cluster investigations are an important public health strategy for responding to public concern about possible associations between disease and environmental exposures (California Department of Health Sciences 1989; CDC 1990; Fiore et al 1990; Neutra 1990).

Public health agencies need to recognise the social dimensions of a cluster, how the community perceives risks, and the influence of the media on that perception. From a public health perspective, a community's perception of a cluster may be more important than establishing the scientific existence of the cluster. The general public may not be satisfied with epidemiological or statistical arguments that deny the existence or importance of a cluster. Achieving rapport with a concerned community is critical to managing the situation.

These guidelines are intended to assist public health officers to respond more effectively and in a timely manner to local community concerns about environmental health issues. They also give a clear plan of action that can be outlined to the first person to report an alleged cluster and, when necessary, to the public and the media.

# What is a cluster?

The term 'cluster' has been used to describe an aggregation of some relatively uncommon disease or event (Last 1988).

The initial characteristics of a cluster are that:

- there is a definable health event
- the situation is generally unusual or unexpected
- there are usually at least two cases of the health event
- there is a perceived closeness of the cases within a time period and/or area defined by the informant
- a potential exposure is suspected, along with an alleged connection between the exposure and the health event
- the informant or the community requests some explanation of the health event.

Three categories of clusters may be reported:

- time clusters – when an unusual number of cases of a disease occurs within a defined time period
- space clusters – when an unusual number of cases of a disease occurs within a defined area
- time–space clusters – when an unusual number of cases of a disease occurs within a defined time period and area.

## Examples of non-communicable disease clusters

Investigations of non-communicable disease clusters reported in the scientific literature cover a broad range of health events, both acute and chronic. A state-wide survey of cluster investigation requests in the United States from 2000 to 2004 found most of the requested investigations were for cancer, followed by birth defects. Public requests were the main impetus for carrying out investigations (Juzych et al 2007). A similar survey of cancer cluster investigation requests found 65 percent of investigation requests came directly from the public. Of those inquiries from the public, 75 percent were resolved at first contact (Trumbo 2000).

In the United States since 2002, the CDC has run a centralised triage system for cancer cluster public inquiries. The response ranges from consulting with state health department staff to participating in epidemiological or biological sampling studies.

Each year the health department in Queensland,<sup>1</sup> Australia responds to about 20 inquiries about suspected clusters, most of which concern cancer (Queensland Health 2012). The number of inquiries in New Zealand is unknown but considered to be lower than in Queensland.

Clusters are common in large populations. From a statistical perspective it is almost inevitable that non-communicable disease clusters will occur in some schools, neighbourhoods or workplaces (Kingsley et al 2007).

Chance is the most frequent explanation of clusters (Neutra 1990). However, many carcinogens have been identified because of occupational or medical clusters. Examples where evaluation of clusters has identified important causal relationships include birth defects and thalidomide, vaginal adenocarcinoma among young women who had *in utero* exposure to diethylstilboestrol, angiosarcoma of the liver and vinyl chloride exposure, male infertility and exposure to the pesticide dibromochloropropane, eosinophilia-myalgia syndrome and L-tryptophan (McBride 1961; Lenz 1962; Herbst et al 1971; Creech and Johnson 1974; Whorton et al 1977; CDC 1989). A literature search by Neutra (1990) found only one neighbourhood cancer cluster investigation that identified a carcinogen. This investigation discovered an association between exposure to the mineral erionite and mesothelioma in a Turkish village. Many other reported clusters have had no obvious common aetiology or have been shown on further investigation to have no more cases than would be expected in the general population.

Some clusters have led to extensive investigation without identification of an environmental cause. In the United States, a local health provider notified state health officials of an increase in childhood leukaemia cases in Fallon, Nevada. Sixteen cases were diagnosed between 1997 and 2002.<sup>2</sup> After a state investigation confirmed a higher incidence of leukaemia, it eventually led to a CDC case-control study of children and their families, biological sampling (urine, blood, cheek swab) for chemicals, viral markers and genetic analysis, and environmental (water, air, soil, dust) investigations. The scope of investigation exceeded that for any previous study of childhood leukaemia. Given multiple comparisons, some findings were expected to be statistically significant due to chance so results of the many data outcomes were reviewed for biological plausibility. In addition, external peer review and community-based panels were used to review the findings, which enhanced communication with the community and affected families (Rubin et al 2007). No environmental cause was identified.

<sup>1</sup> The population of Queensland is slightly higher than New Zealand's population.

<sup>2</sup> Fewer than two cases were expected.



Extensive investigation of four cases of a rare birth defect, sirenomelia (or mermaid syndrome), born at one Colombian city hospital in a 55-day period similarly did not identify an environmental cause, although a neighbouring landfill could not be definitely excluded (Orioli et al 2009).

Occasionally extensive investigations are needed because of persistent occupational health and safety concerns.

In 2002 an alleged cancer cluster among security staff at the National Gallery of Australia was identified from a review of sick leave; an investigation found no unusual occurrence of cancer. However, due to ongoing concerns, including about the initial investigation, a detailed investigation of past and current exposures to carcinogens in the Gallery and cancer incidence in current and former employees was initiated in 2006. While excess colorectal cancer was found in security officers, there was nothing to suggest it was related to work-related exposures (Driscoll et al 2008).

In 2005 concerned staff at the Australian Broadcasting Corporation (ABC) studios, Brisbane reported an alleged cluster of breast cancer to health authorities. The number of cases and young age profile led to a decision to investigate. Given there was no specific environmental exposure, it was difficult to identify and quantify the exposed population. Different assumptions about workforce size were made but a statistically valid conclusion was not possible. This led to an environmental investigation to identify possible exposures. However, environmental testing was not conducted as evidence of potential exposure pathways to known breast carcinogens was insufficient.

Neither the health department investigation nor a preceding investigation by an occupational health physician of cases and survey of radiofrequency electromagnetic radiation allayed concerns. The women expected environmental testing and in 2006, after another case was diagnosed,<sup>3</sup> the ABC appointed an independent review and scientific investigation panel. The investigation addressed community needs by altering its direction and content to include environmental factors that, although unlikely to be causative, were of concern to the staff and public (Stewart 2007). The panel investigated for known or suspected exposures that cause any cancer, not just breast cancer, which might plausibly be present. The age-adjusted risk was related to duration of employment and the panel concluded some aspect of work or the work environment may have contributed to it (Armstrong et al 2007). Staff concerns were finally allayed by relocation from the site, even though they recognised that no specific cause of the cluster was found.

In case an unknown or undetected exposure was present in other ABC studios, female employee records were linked to the cancer registry, complemented by self-report. Excess breast cancer was not found in all ABC female staff, including or excluding Queensland, compared with respective general population incidences (Sitas et al 2010).

<sup>3</sup> In total, 13 cases were identified from 1994 to mid-2006.

# Clusters in New Zealand

Few reports of clusters in New Zealand have been published other than in the media. Media reports of clusters are often inaccurate and may be extensive, leading to a public perception that clusters are common and reinforcing concerns about environmental exposures.

In the 1970s a number of alleged clusters of spina bifida linked to the herbicide 2,4,5-T were reported (Sare and Forbes 1972; Department of Health 1977). A Department of Health inquiry into clusters in Taranaki, Northland and Waikato found no excess incidence of spina bifida and no evidence to implicate 2,4,5-T as a causal agent (Department of Health 1977).

In 1990 the Department of Health investigated a reported cluster of congenital cataracts in the Wellington area. No excess incidence and no common aetiology were demonstrated (Elwood 1990). In 1993 an alleged cluster of birth defects in the children of three former Christchurch City Council horticultural workers, which were said to be linked to exposure to a fungicide, benomyl, received intense media scrutiny (Borman and Read 1995). Two of these children had eye defects. An independent inquiry found no unusual occurrence of these birth defects (Alchin 1994).

A study of the health status of Auckland fire fighters involved in a major chemical fire found a cluster of testicular cancer among the comparison group of Wellington fire fighters (Bandaranayake et al 1993). Investigation of the cluster confirmed an excess incidence of testicular cancer among Wellington fire fighters in the 1980s but no causal factor was identified and chance could not be excluded (Bates and Lane 1995).

Clusters of suicide in prisons and police cells and in the community, as well as of attempted suicide among young people, have been identified in New Zealand from analysis of mortality and hospitalisation data (Cox and Skegg 1993; Gould et al 1994; Larkin and Beautrais 2012). Analysis of hospitalisation data has also found urban time–space clusters of childhood asthma (Hales et al 2005). Detection of non-communicable disease clusters by scanning data for evidence of excess risk requires subsequent further study to investigate causal factors.

In 2004 a cluster of thyrotoxicosis was confirmed following notification by an endocrinologist of four cases seen over one month from the same Otago area. Investigation identified one further case seen at a similar time in another area. All cases had regularly consumed the same brand and flavour of soy milk which had, independently of the cluster investigation, been found to have elevated iodine (from kelp added for flavour) and been reformulated. A case-control study confirmed the association with soy milk consumption (O'Connell et al 2005).

In Nelson a group of women whose husbands died of motor neurone disease (MND) raised concerns of a possible link between six cases of MND among Port Nelson workers and a fumigant, methyl bromide. Investigation found the number of cases diagnosed in the Nelson/Tasman area from 1995 to 2005 was consistent with the expected incidence. No excess MND mortality in the Nelson/Tasman area was found compared with New Zealand as a whole or other regions where methyl bromide use was greater. The incidence among port workers could not be calculated due to lack of denominator population data for the Port Nelson area. The investigation concluded that MND in this group of workers was most likely due to chance. Three cases could have had some exposure to methyl bromide as a result of proximity to timber fumigation. Methyl bromide is not a known risk factor for MND and no evidence linking MND and methyl bromide was found in this investigation (Kiddle 2005).

## Surveillance for clusters

Active surveillance to identify clusters is best carried out in the workplace where the population at risk and the exposures are limited and can be defined. Routine analysis of registries or vital statistics by public health agencies for unsuspected clusters has not been recommended as a valid public health exercise (Smith and Neutra 1993; Elliott et al 1995). Smith and Neutra (1993) argue against such a recommendation on the grounds that:

- cluster reports from the public still require a response
- most diseases do not have a timely, up-to-date registry
- information routinely collected by registries is often inadequate for investigating clusters (eg, address at the time of diagnosis may not be aetiologically relevant)
- analysis tends to focus on geographical clusters and may not identify occupational or other clusters
- routine analysis identifies false positives as well as false negatives. The cost of a false positive to a community may be considerable in terms of outrage, concern about personal health, loss of business and/or falling property values.

Smith and Neutra (1993) advocate a more appropriate surveillance activity is for public health agencies to be vigilant in detecting new and unusual exposures and to evaluate their impact. However, in the United States, registries for chronic disease surveillance have increasingly been established and used. This trend has been accompanied by the development of guidance about use of GIS techniques and cluster analysis software (Kingsley et al 2007).

## The Texas sharp-shooter fallacy

Looking for clusters is analogous to the Texas sharp-shooter fallacy. A sharp-shooter fires randomly at a wall and then draws a bull's eye around the area where most of the bullet holes are, making it look as though the accuracy has been excellent. This definition after the event also applies to boundaries in time and person as well as space.

Some clustering may be expected to occur by chance due to random distributions. The greater the number of possible cases of a disease and of areas and time periods that are examined as potential clusters, the greater the chance that randomly distributed cases will appear as a cluster. For example, if the rate of a disease in 100 areas is looked at and cases are occurring randomly, you would expect to find a statistically significant increase in about five areas.

## Occupational clusters

Occupational clusters are likely to continue to contribute to the understanding of relationships between disease and exposure that are not predicted by toxicological studies at the time when new chemicals and work processes are developed.

Advantages of the workplace as a setting for cluster investigation include natural denominators as a result of shared work areas and identifiable time periods associated with work processes, shared exposure(s), the ability to form hypotheses based on job descriptions, and the possibility of finding comparable populations in which these hypotheses can be tested (Fleming et al 1991).

Methods have been developed for the initial assessment of occupational cancer clusters using limited data such as basic data about the cluster and the size of the workforce, and the number of workers entering and leaving the workforce in each year (Smith et al 1994). Fleming et al (1992) have developed a step-by-step protocol for the investigation of disease clusters in the workplace.

## Cluster investigation and causation

If a cluster is found determining causation is difficult. Characteristics that increase the likelihood of detecting a causal agent through a cluster investigation are that:

- there are at least five cases and a very high relative risk (RR) (eg,  $RR \geq 20$ )
- a unique and detectable class of agents has been responsible for the disease in the past
- the pathophysiological mechanism for the disease is well understood

- the agent persists in the environment and can be measured there
- the agent persists or leaves a physiological response in exposed individuals and is rare in the normal population
- exposure is heterogeneous within the community
- accurate self-exposure assessment is possible by questionnaire or can be obtained from records
- a multi-community study consisting of some similarly exposed and unexposed communities is feasible
- the cluster represents an uninvestigated endemic space cluster rather than a time–space cluster. This scenario suggests a stable, persistent problem and possibly a persistent agent (Neutra 1990).

Bradford Hill's nine criteria,<sup>4</sup> or guidelines, provide a framework against which the causal nature of an association between an exposure and outcome can be assessed (Lucas and McMichael 2005).

<sup>4</sup> Bradford Hill's criteria are strength of association, consistency, specificity, temporality, biological gradient, plausibility, coherence, experimental evidence and analogy.

# The cluster investigation process

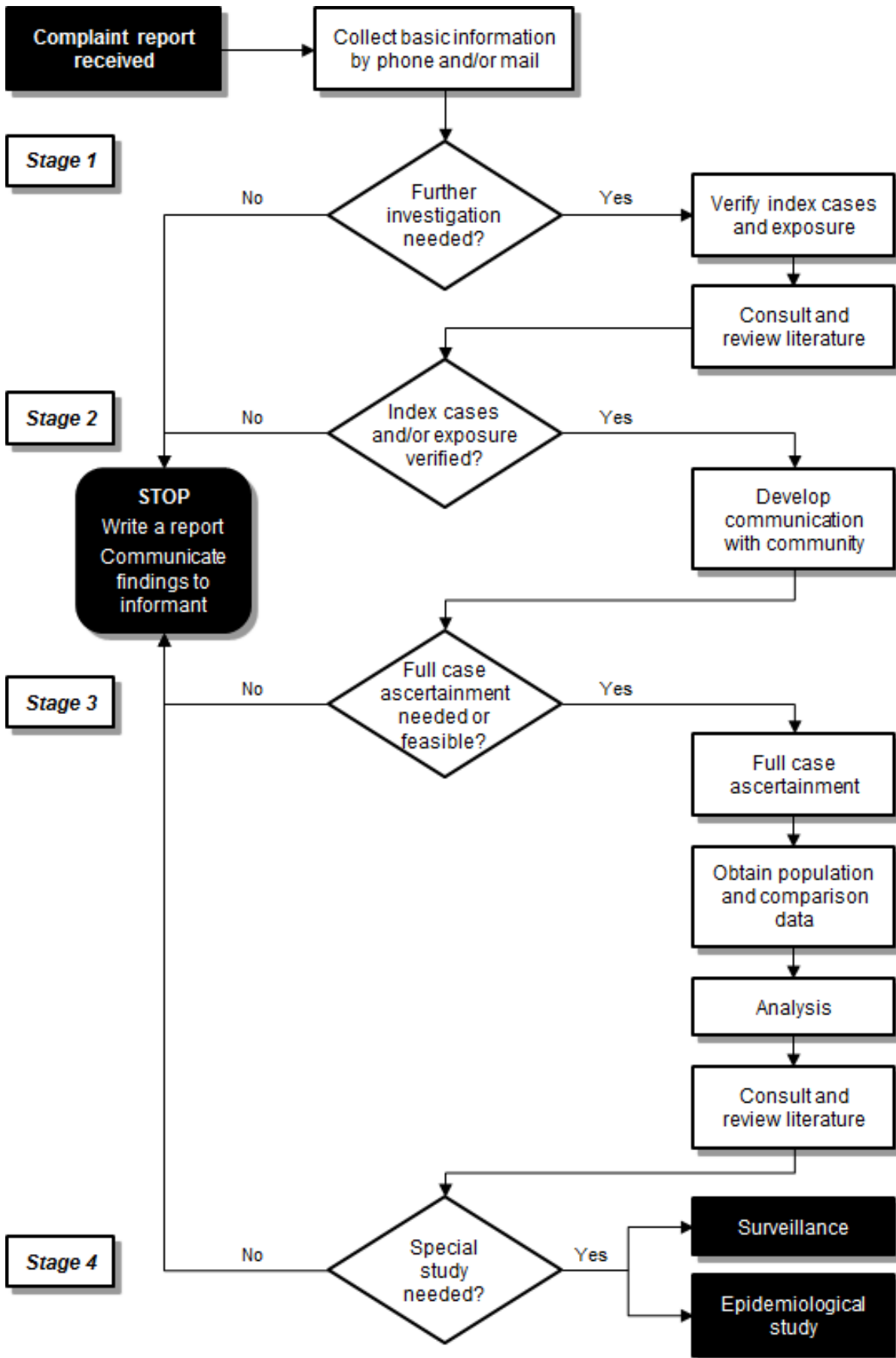
The investigation of a cluster of non-communicable disease has four distinct stages (Figure 1). Each successive stage involves collecting more specific, but more varied data and verifying those data more strongly. The boundaries between these stages are flexible. Depending on local judgement, experience and the available resources, a public health officer may choose to combine stages.

At the end of each stage a decision must be made about whether to proceed further. How to communicate the results of that decision to the public and other interested parties must also be determined.

This step-by-step approach follows the principles of epidemiological research, which are to:

- establish that a problem exists
- confirm the homogeneity of the events
- collect data on all events
- characterise the events in terms of epidemiological factors
- look for patterns and trends
- formulate a hypothesis
- test the hypothesis
- write a report, obtain peer review and communicate the results.

Figure 1: Flowchart of the overall cluster investigation process



# Stage 1: Preliminary evaluation of a report of an alleged cluster

The first stage is the response when an alleged cluster is initially reported to a public health officer (Figure 2). If a cluster is suspected from monitoring or vital statistics, the procedure begins at Stage 2.

## Step 1: Record the initial report

Alleged clusters can be identified by anyone, including members of the public, media, health professionals, or local, regional or national agencies, or through environmental or health monitoring systems and vital statistics.

The public have a serious, quick and often adverse reaction to any thought of a cluster. They want the public health authorities to treat the matter seriously and with concern. The successful management of public outrage depends to a considerable extent on how the initial report is handled.

Whoever receives the report should identify himself or herself to the informant and tell the informant what actions will be taken, how long these will take and when a response can be expected.

To avoid overlooking vital data, it is advisable to use a standardised form (see the appendix) to collect this initial information. The informant's real concern may also only emerge in response to careful questioning.

When an alleged cluster is initially reported, get as much information as possible about the informant and the index case(s). This can save time and resources later and also indicates that the report is being treated seriously.

## Step 2: Form an initial case definition

The initial case definition is based on the following questions.

- What is the specific disease, symptom or health event of concern?
- Where is the affected geographical area, population group or workplace?<sup>5</sup>
- When did the specific disease, symptom or health event occur?
- Who (eg, age, ethnicity, sex) are the index cases (ie, the cases first reported)?
- What, if any, are the suspected specific exposures?

<sup>5</sup> Refer an informant with workplace concerns to WorkSafe New Zealand.



Consultation with surveillance and clinical experts early in the process can help ensure the case definition is appropriate.

The case definition for birth defects identifies:

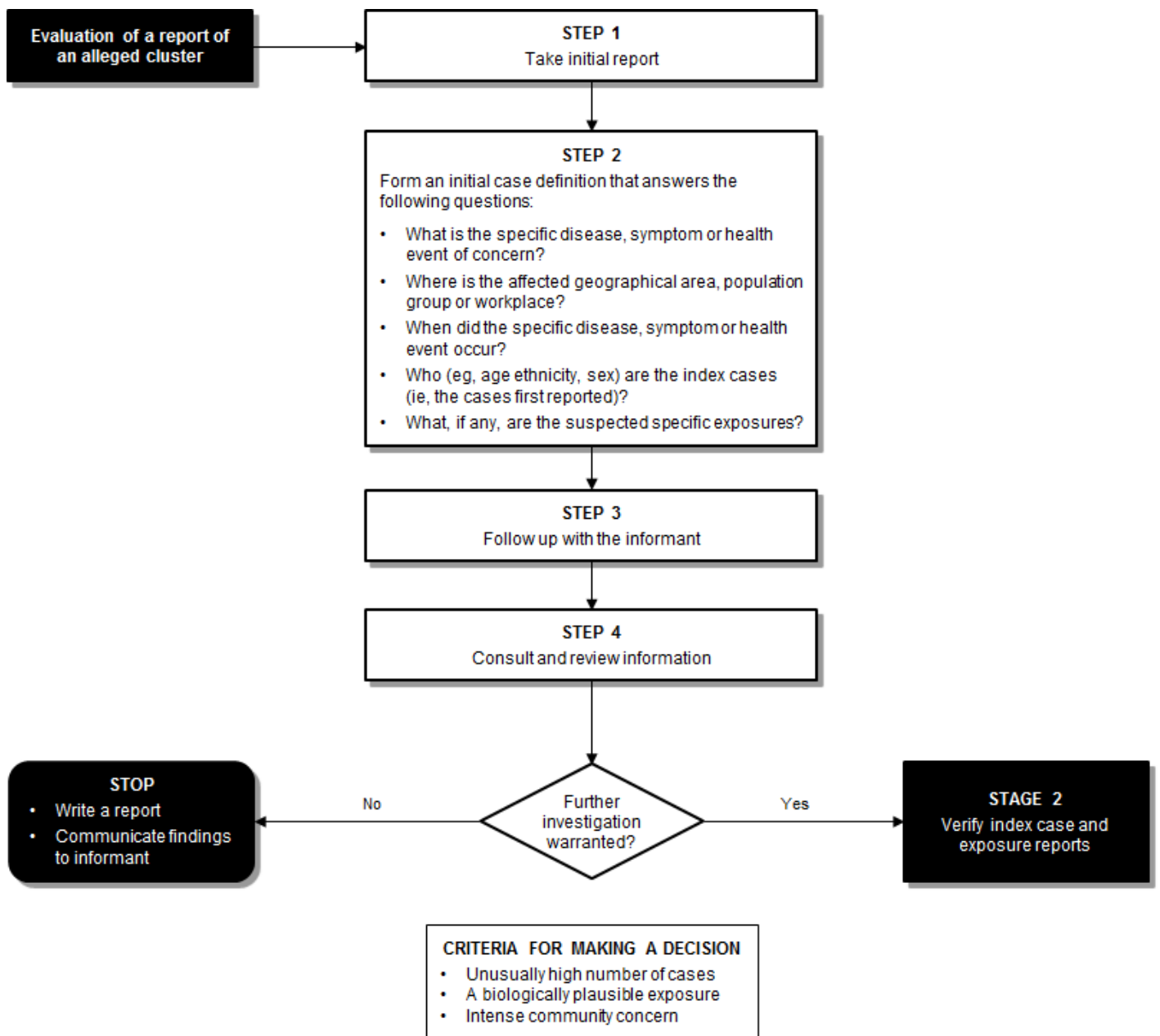
- the type of defect(s) (the same or pathogenetically similar defects)
- the time period based on the defect when a potential exposure could have a causal role
- the estimated time of conception
- the age at which the defect is diagnosed (eg, before one year)
- whether prenatally diagnosed cases are included.

For some defects, the methods of diagnosis may be included in the case definition. For defects with unknown aetiology, the time period may include the time from one month before pregnancy through pregnancy. For other defects, a narrow time period exists (Williams et al 2002). For example, the neural tube defects anencephaly and spina bifida occur within 28 days of conception. Potential exposures must therefore have occurred before this 28-day maximum.

A site visit may be warranted to gather general information about local exposure possibilities. Often the public focus on a source (eg, landfill, contaminated site or industrial facility) rather than one or more specific exposures that might have occurred. The site visit aims to answer the following questions.

- What hazards are present?
- What is the geographical location of the hazards in relation to the population at risk?
- Are there known or potential exposure pathways by which these hazards might have affected the population at risk?

**Figure 2: Flowchart for Stage 1 – Preliminary evaluation of a report of an alleged cluster**



## Step 3: Follow up with the informant

Many reports of an alleged cluster can be resolved at this step by an explanatory letter, email or telephone call. The investigation of clusters should have a strong health education component and community involvement (Neutra 1990). Anxiety about a cluster can often be lessened by telling the informant that:

- a disease that the public perceives to be rare occurs quite often. For example, cancer is a relatively common disease, most types occurring at a rate of about one per 100,000 person-years, and the risk increases with age; major birth defects occur in 1–2 percent of live births
- the length of time that cases live in the cluster area must be substantial to implicate a plausible environmental carcinogen, because there is a long latency for most known carcinogens
- cases that occurred among people who are now deceased may not be helpful in linking exposure to disease because of the lack of data on exposure and on possible confounding factors
- the occurrence of clusters of specific diseases in a population is often due to chance.

There will rarely be a single explanation for the occurrence of a cluster. Many cluster investigations of non-communicable disease are initiated because one factor, often an easily identified environmental hazard, is suspected as the cause. However, most diseases have a number of possible causal factors, involving an interaction between the individual and the environment. For example, even though thalidomide is one of the most powerful teratogens, some pregnant women who took it did not give birth to an infant with birth defects; the exposure had to occur during a specific time period in the embryological period.

## Step 4: Consult and review information; make a decision

The decision as to whether to further investigate the alleged cluster is made after reviewing the initial information and consulting with specialists in cluster investigations and in the disease(s) concerned. Let the informant know the outcome of the decision as soon as possible.

In general, further investigation is warranted if any one of the following conditions exists:

- an unusually high number of cases
- a biologically plausible exposure(s)
- intense community concern.

In one American state, the criteria for undertaking a preliminary investigation are:

- there are at least three cases in a defined population (two cases for very rare diseases)
- health effects are related, through physiology or exposure pathway
- cases lived in the area during the exposure period and not merely at the time of diagnosis (Montana Department of Public Health and Human Services 2006).

Often investigation must go beyond what is indicated scientifically in order to address community concern.

If it is decided that the alleged cluster demands further investigation, the informant could become an active participant in the data collection relating to the cluster. A member of the public may have easier access to some data that would be helpful to the investigation.

If a decision is made to end the investigation, the following actions are recommended:

- write a report with a summary and conclusion
- obtain appropriate peer review
- communicate the results to the public. This could involve writing a letter or email to the informant, or a public announcement, including media releases and public meetings. A written response to the person or organisation who initially reported the cluster will reduce the possibility of any misunderstanding about what was done and when.

Many reports of alleged clusters arise from genuine concerns among the public about either their current or future health. They want action, answers and reassurance. All reports need to be treated seriously until there is evidence to the contrary. Public concerns can often be allayed by prompt action, keeping the public informed and openly providing as much information as possible.

Preliminary communication involves setting limits for the investigation including agreement on terms of reference, the disease, area and time period to be studied. This is very important in terms of achieving community satisfaction with the process (Gavin and Catney 2006). It is also important to indicate how long the investigation may take and its limitations, such as the difficulties of establishing a causal link between the environment and disease.

## Stage 2: Verification of index case and exposure reports

The next stage in the investigation is to verify the index case(s) and suspected exposure(s) that have been reported (Figure 3). If not already undertaken, a site visit is advisable. In many instances the resources required to undertake an extensive verification are not available locally, and assistance from other specialists and agencies is needed.

As many alleged clusters are associated with high community concern, careful documentation of telephone conversations and meetings helps avoid later disputes.

### Step 1: Establish who should do the verification

To verify the index case(s), it is advisable to consult with appropriate specialists (eg, a pathologist, neurologist, clinical geneticist, toxicologist) or specific agencies (eg, Te Whatu Ora, WorkSafe New Zealand). An environmental health expert or an occupational hygienist may be needed to verify the possibility of exposure. Often a perceived exposure is not verifiable (Fiore et al 1990). Measurements of exposure are usually unnecessary at this point.

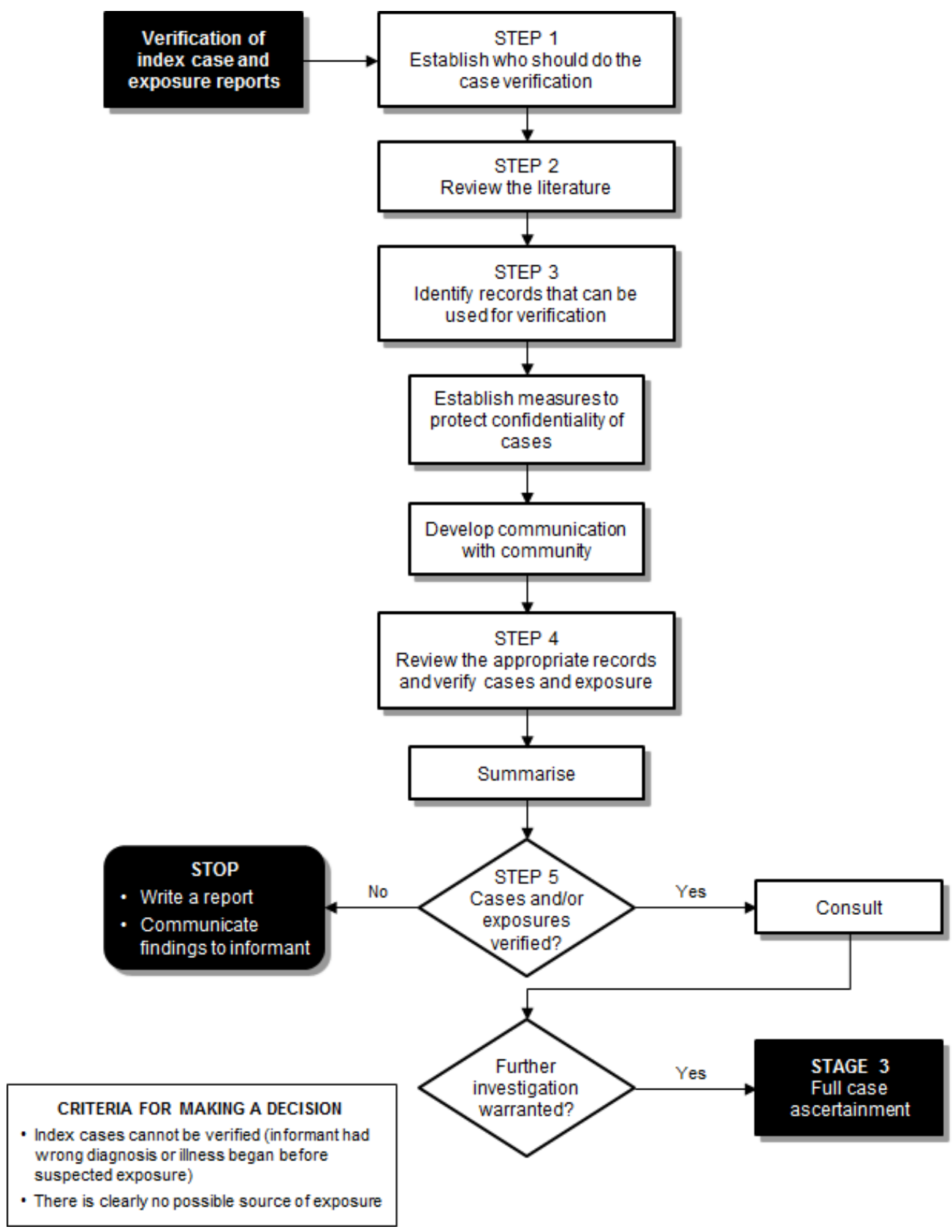
### Step 2: Review the literature

The investigator reviews the literature for evidence of any previously reported clusters of the disease, known or hypothesised exposure associations and other epidemiological and toxicological information. Online databases such as Hazardous Substances Data Bank<sup>6</sup> may be useful when exposure to specific hazardous substances has occurred or is suspected.

The purpose of the literature review is to determine what is already known or hypothesised about potential causes of the disease and alleged exposure(s).

<sup>6</sup> <https://www.nlm.nih.gov/toxnet/index.html> (accessed 6 September 2023).

**Figure 3: Flowchart for Stage 2 – Verification of index case and exposure reports**



## Step 3: Identify the records that can be used for verification

Records that can be used to verify cases are:

- death certificates, including infant death certificates
- medical records
- population-based registries (eg, New Zealand Cancer Registry,<sup>7</sup> New Zealand Congenital Anomalies Registry<sup>8</sup>)
- employment records.

Exposure is relatively easy to determine in acute disease clusters and can be done by questionnaire. In most instances, exposure has been through personal contact or through food, drug or beverage consumption. In contrast, exposure through water, air, soil or dust is poorly correlated with questionnaire responses (Neutra 1990).

Potential exposure(s) can be verified from: agency and company files about sites and facilities in the area of the cluster or where the index cases lived or worked; aerial photographs; records of water, soil and air quality from various monitoring agencies; and planning records about previous industrial sites and property uses.

Availability and access to information may be constrained by the Privacy Act 2020 (see the next section, Legislation). The informed consent of the index case(s) (or their next of kin if the index case is dead) may be necessary to access records.

## Step 4: Review the appropriate records and verify cases and exposure

In many instances, the easiest way to verify a disease is to review hospital records, in particular any diagnoses on pathology reports. Diseases can also be verified from doctors' records and the cancer and birth defects registries. Cause of death can be confirmed by using the death certificate and reviewing appropriate medical records. The case definition provides guidance in deciding what data are to be used.

Verification of exposure is often more difficult because of the comparative lack of relevant data. It may be helpful to consult with epidemiologists and occupational and environmental health experts to gain information about the availability of, access to and use of data such as employment records and residential histories.

<sup>7</sup> <https://www.tewhatauora.govt.nz/for-the-health-sector/health-sector-guidance/diseases-and-conditions/cancer/new-zealand-cancer-registry-nzcr>

<sup>8</sup> <http://www.ehinz.ac.nz/projects/new-zealand-congenital-anomalies-registry/>

GIS may be used to visually display the alleged cluster and exposure source(s) and explore spatial patterns of potential exposures.

Early in the investigation of a cluster, there may be requests for new environmental data to be collected. Premature environmental measurements should be avoided because they may be unfocused and uninterpretable.

## **Limitations of records**

Many of the available data will have been collected for purposes other than for investigating a cluster. As a result, the recording of information may vary between sources and make the information difficult to interpret. For example, the hospital discharge form's diagnosis might be pneumonia with no indication that the condition may be the result of exposure to a toxic substance. An individual's cause of death also may not be identical to the reason for hospitalisation (eg, a subsequent pathology report could indicate a previously unreported cancer).

Many studies have shown that the level of disease ascertainment is directly related to the number of records used in searching for cases. Studies using multiple sources have a higher validity and level of ascertainment than studies involving only a limited number of sources.

## **Step 5: Make a decision – stop and report or investigate further?**

At this stage of the investigation, it is helpful to summarise the findings to date in writing and review them. This may stimulate new ideas about the disease or possible causes.

Challenges particular to clusters of birth defects include difficulties in case classification, problems in case ascertainment (lack of data on stillbirths and terminations, the time period of ascertainment, the perinatal autopsy rate, the use of single or multiple data sources) and limited information on aetiology (Williams et al 2002).

Suicide clusters differ from other clusters because the community's perception that a cluster exists may itself be an important risk factor for further suicides and attempted suicides. Action is required, regardless of the number of cases reported. A community response should be initiated to identify individuals at high risk of suicide and refer them for assistance (O'Carroll and Mercy 1990).

Reported clusters will fall into one of three categories.

### **1 No cluster**

The initial investigation will often find the number of cases is not excessive. In many instances, the original disease and/or exposure allegations are not supported by the



medical records and/or environmental inspection, or else the examination of records does not confirm the suspected environmental exposure.

The possible effects of migration are also important. Some of the cases involved in the reported cluster may have developed the disease before moving into the area and encountering the possible exposure, and they should not have been included.

The disease or exposures alleged in the reported cluster may be a number of different diseases or exposures. The term 'birth defects' includes a wide range of specific defects that have different epidemiological characteristics and are likely to have different aetiologies. When the diagnoses of the reported cancer cases have been verified, they may be different types of cancer or may not be cancer. It is unlikely that unrelated cancers will constitute a cluster.

## **2 Explained cluster**

Many reports about clusters of cancer, spontaneous abortion, and birth defects arise because the public do not realise how common these conditions are. For example, after a clear explanation, the public are likely to understand that a few cases of lung cancer in a retirement community with a high percentage of smokers and no unusual environmental exposure are unlikely to constitute a cluster. A high number of Down syndrome births in a local population with a high proportion of older mothers is also not uncommon.

The investigation can be stopped if the reported cluster fits in either of the above categories. In this case, a written report should be produced, summarising the investigation to date and justifying the decision to stop.

## **3 Unexplained cluster**

If the reported disease and/or exposure are confirmed, the investigator must decide whether to proceed to Stage 3 (full case ascertainment). This decision depends on the type of disease(s) and exposure(s), the size of the apparent cluster and the biological plausibility of a disease–exposure relationship.

Some of the factors that support biological plausibility are that:

- the exposure is capable of causing the disease of concern
- the exposure is of sufficient magnitude to cause adverse effects
- all cases have been exposed
- the temporal relationship between the exposure and the disease is in keeping with what is known about the disease (eg, latency periods) (ATSDR 2002).

To prevent any misinformation or confusion, it is preferable to let the original informant know the decision by both a:

- personal telephone call, which gives them the chance to ask questions
- follow-up letter or email, to document the information clearly.

## Stage 3: Full case ascertainment

This stage involves finding and verifying all additional unreported cases of the disease(s) in question in the time period and geographical area of interest (Figure 4).

### Step 1: Establish a case-finding team

It is often helpful to form a case-finding team because of the complexity of the environmental health issues and the need for a variety of disciplines and skills to collect and analyse the data. The team needs to include, as a minimum, specialists in epidemiology, environmental health (if exposure verification is needed) and public health. The roles and responsibilities of team members, the channels for communication between members and the spokesperson to the media or public should be decided at the outset.

### Step 2: Revise the case definition

It is crucial at this stage of the investigation to review and if necessary revise the initial case definition. A complete case definition includes a:

- definition of the health events to be counted
- time period during which diagnosis occurred
- geographical area and/or population group of interest.

#### 1 Health events

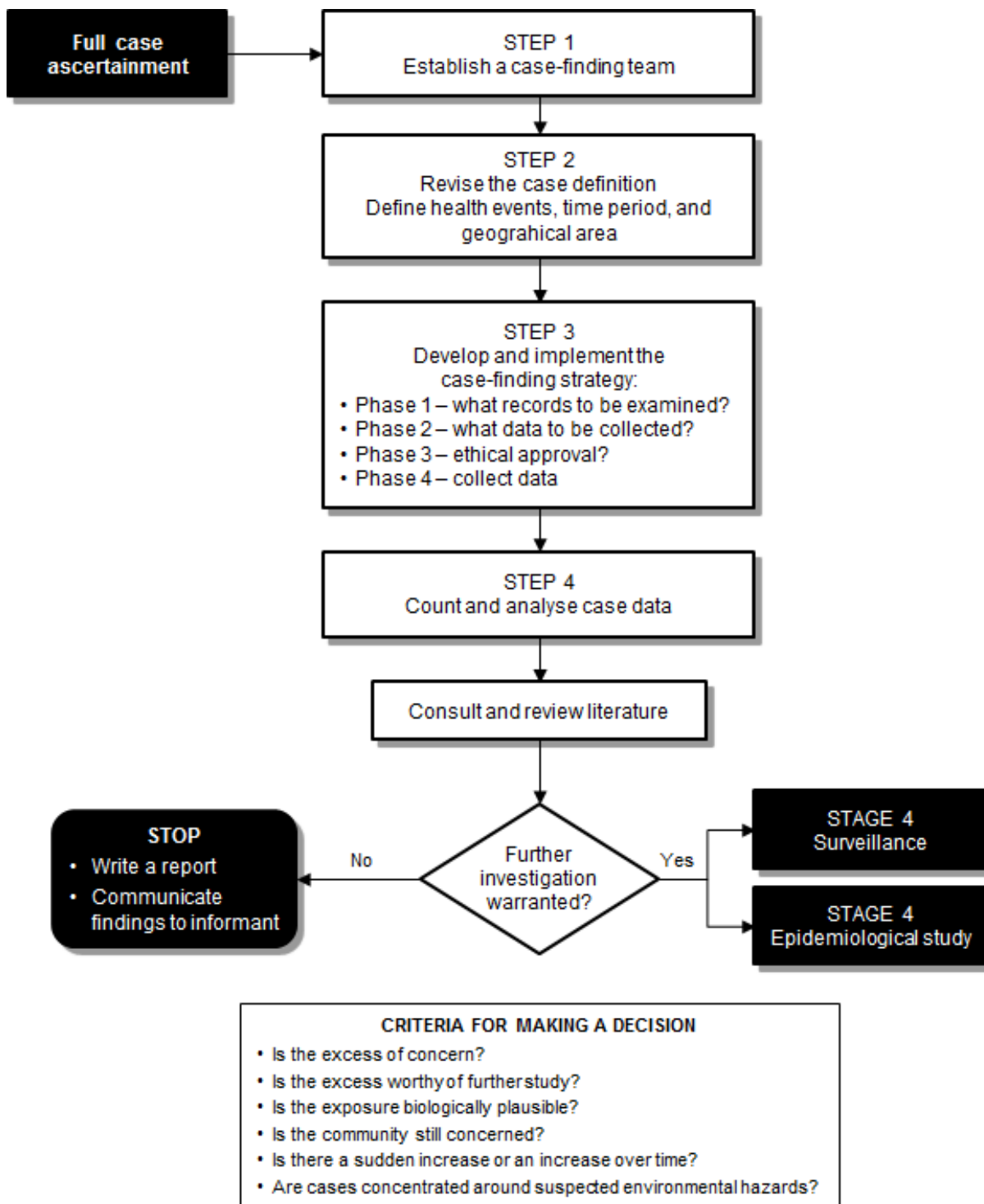
Only the specific disease or closely related diseases suspected of clustering are counted, as in the following examples.

- If squamous cell carcinoma of the lung is reported, all types of primary lung cancer would be counted because they might be caused by the same type of exposure. Other types of tumours would not be counted.
- For birth defects, both live births and stillbirths should be included and, if possible, other types of reproductive outcomes (eg, spontaneous abortions). If the investigation initially focused on spina bifida, all types of neural tube defects should be included. Many birth defects (eg, anencephaly) have a higher rate in stillbirths than in live births. If the investigation does not collect stillbirth data, it will exclude an important number of cases. It is also important to decide on the follow-up period in clusters of birth defects. A number of birth defects (eg, heart defects) are not diagnosed until after the first year of life, and there may have to be an extended follow-up period.

A broader set of conditions should be counted if there is concern about general increases in diseases that might have resulted from exposure to a mixture of toxic chemicals, or if a number of unrelated diseases were reported.

Cases that cannot be confirmed using medical records should be tabulated separately or not be counted.

**Figure 4: Flowchart for Stage 3 – Full case ascertainment**



## 2 Time period

The best time reference point for a chronic disease is the calendar year of diagnosis, but occasionally only the date of death is known. A time period of possible exposure needs to be defined and all cases diagnosed during that period need to be identified. Clusters of cancer also need to consider an appropriate latency period. As the correct latency period is often unknown, a range of latencies (eg, 0, 5 and 10 years) is assumed based on the number of cases.

When investigating clusters of birth defects, focus on possible exposures at or before the time of conception, rather than at the time of birth. Exposures at the time of birth will not necessarily be the same or at the same level as they were at the time of conception. Most birth defects have a critical period when the defect can occur. For example, neural tube defects occur within the first 28 days after conception, so the occurrence of these defects will not be due to any potential exposure after that time.

## 3 Geographical area or population group

The community of interest (eg, suburb, city, health district or territorial authority) is the basic areal unit for data collection and analysis involving possible community exposure(s). It is often more meaningful to disaggregate heavily populated areas into smaller areas. If the focus is on the occurrence of disease among a specific population subgroup (eg, a particular neighbourhood or all women over the age of 35 years), the smallest unit that includes the entire group and for which statistical data are available should be used as the denominator. This unit is also used for collecting information about membership in the subgroup. Alleged clusters are unlikely to correspond to the administrative boundaries that have determined reporting of population data or health statistics (Elliott 1995).

If the focus of the investigation is potential exposure at the workplace, the work site itself is usually the basic unit of analysis. Specific occupations or subgroups working in high-exposure areas may be defined, but counting cases in the entire workplace provides the basis for comparisons between subgroups.

Whatever data are collected for the cases should also be collected for the reference or unexposed population. The coding of data should also be consistent within and between these two groups.

Arbitrary judgements have to be made about the selection of the reference population. Often selection is constrained by the coverage and extent of available routine data (Coory and Jordan 2013).

Domicile data should relate to the place of usual residence. This may not be the domicile at the time of diagnosis. Data on the length of residence in a particular area are also important, especially for diseases such as cancers, in which exposure may have occurred 10 or more years previously. Without these additional data, a person who has lived one

year in an area is assumed to have a similar degree of exposure to a person who has lived in the same area for 10 or 20 years. In reality, the person with a short residency in an area may have been exposed to a possible environmental hazard in another area, or may not have been exposed to an environmental hazard in the current area.

In the New Zealand mortality statistics, domicile is given as the usual place of residence. There is no definition of what constitutes 'usual', so this can range from months to years or life. The address of a family member may also be given as the usual place of residence even though the case has only temporarily stayed there.

In clusters of birth defects, it is important to know the domicile of the mother at the time of conception and, if possible, during the year before the birth. For childhood cancers, attempt to establish possible exposures from the time of gestation. In occupational clusters, the full work history – not only the most recent occupation – is relevant. Occupation is recorded in the national mortality statistics as the occupation at the time of death.

The following examples illustrate how the above guidance can be applied in case definitions. In a cluster involving a pesticide-induced food illness, the case definition might include all individuals who experienced vomiting or diarrhoea within two hours of eating produce anywhere in a defined geographical area. The initial case definition in a possible childhood cancer cluster might include all cancers that occurred during the last five years in children younger than 15 years who were living in that geographical area before they were diagnosed.

## Step 3: Develop and implement the case-finding strategy

The process of finding the cases and collecting the data involves four phases.

**Phase 1: Decide what records need to be examined.** All cases with the disease(s) that were diagnosed in the area or workplace and time period need to be identified. This usually involves finding and reviewing data from several sources, including:

- medical records
- death records
- population-based registries (eg, New Zealand Cancer Registry, New Zealand Congenital Anomalies Registry)
- employment records.

Case finding could also involve laboratories, pharmacies, disease societies (eg, Motor Neurone Disease Association) and direct appeals to general practitioners, certain physicians (eg, neurologists) and the public.

Each data source has its own particular strengths and weaknesses. For example, hospital records are a good source of information, but may include unconfirmed cases, such as suspected diagnoses, and omit cases not diagnosed in the hospital or that were diagnosed in another hospital. Death records may have a vague, non-specific diagnosis or may omit a diagnosis when it was not the underlying cause of death.

Disease registries are a good source of cases for full case ascertainment. However, they may not be available in the area for the time period of interest. In the absence of a registry, full case ascertainment is more difficult and resource-intensive.

**Phase 2: Determine what data will be collected.** Data availability depends on whether the source of information is registries, medical records, or interviews. Use a standardised, structured questionnaire. In general, the following are the minimum data collected for each case:

- name or other identifier
- date of birth
- ethnicity
- sex
- age at time of diagnosis and/or death
- residence at the time of diagnosis
- diagnosis and basis of diagnosis.

If possible, also collect data on:

- family history of the disease(s) in question
- known exposures, such as smoking
- length of residence at the current address
- history of past residence in the area of interest
- other relevant potential exposures – for example, playing in contaminated fields, drinking contaminated water, eating contaminated foods, overseas travel, hobbies and other activities.

If investigating a possible exposure in the workplace, take a detailed occupational history. This includes:

- occupation of the case or parent
- employment history – job type, classifications and duration at each position – as far back as possible.

The following workplace information is also desirable:

- known work exposures to toxic substances – for example, asbestos exposure of the case or parent
- identification of potential hazards at the work site.

**Phase 3: Obtain ethical approval**, if required, to carry out the data collection.

**Phase 4: Collect the data.** The data collection methods chosen depend on the type of data needed to count all suspected cases. Much of the basic information can usually be obtained by reviewing existing data. It is important to indicate which source document is used to obtain the data so that data validity can be assessed. Data must be recorded and coded in a systematic and consistent manner.

A questionnaire should be designed so that all the necessary data are obtained in a clear and unambiguous manner from one interview. Consultation with experts about the technical aspects of questionnaire design, pretesting, training of interviewers, and data coding and processing is advisable.

It is important to adhere to confidentiality during data collection and storage.

## **Step 4: Count and analyse case data**

Once every effort has been made to find all the cases in the cluster population that conform to the agreed case definition, the data are examined and any duplicate case reports are eliminated.

Numerator data for calculating cluster and background disease rates can be obtained from the national health statistics – for example, mortality, hospital discharges, the New Zealand Cancer Registry and health surveys. Data specification (eg, by age groups and ethnicity) should be in the same format as for the denominator data. Rates are not routinely calculated for small areas such as a neighbourhood because the number of observed cases is usually too small for stable rates or meaningful analysis (Fiore et al 1990).

Population data are necessary to calculate expected numbers of a disease based on published background rates of disease. The expected number of cases can be compared with the actual number of cases observed in the study population to determine whether the community has experienced an excess rate of the disease.

Cluster investigations usually require detailed population data (eg, by sex, ethnicity and age groups) for very small geographical areas. Such data can be obtained every five years from national census data. Data for non-census years can be requested from Statistics New Zealand. Denominator data may contain significant errors in non-census years for areas with small populations (Elliott and Wartenberg 2004). Using census data in a non-census year assumes that there has been no change in the demographic composition of

the population. Denominator data can also be obtained from a number of government departments, including the Ministry of Education - Te Tāhuhu o te Mātauranga.

In epidemiological terms, the number of cases in the study or cluster population is the number of observed cases. The number of expected cases of a disease is determined by multiplying a background rate of disease by the study population in the time period and geographical area that was used in counting the observed cases.

Once a complete (or virtually complete) count of cases has been established, it must be decided whether the number of cases that has been observed is different from the number expected. Usually cluster investigations are concerned with determining if there is an excess number of cases in a local population. A finding of fewer than expected cases is, however, reassuring for the community.

Often a cluster is thought to be 'real' if there is a statistically significant excess of cases. The statistical tests used in the analysis calculate the probability that the disease rates observed in the cluster would occur by chance alone. This will usually involve comparing the observed rate in a known population with expected rates derived from larger population surveys or disease registries.

As most diseases are not evenly distributed throughout a population, an observed excess (or deficit) of cases may, therefore, occur at random such that the cluster is not aetiologically important. There are many census areas and small towns in New Zealand, and hundreds of thousands of workplaces, social groups like clubs, and churches. All of these groupings are at risk of excess rates of a non-communicable disease, even if the distribution of the disease itself is random. The number of observed cases will rarely equal the number expected, even without an environmental cause. The pivotal question is, 'How far away from the expected number must the observed number be to make it a very unusual occurrence?'

Standard statistical and epidemiological techniques for assessing excess risk can often be used to evaluate reported clusters, but statistical significance should not be used as the sole criterion for investigating a disease cluster. A small observed number of cases may be worth investigating if a biologically credible exposure is present or public concern is intense.

A useful first step is often to produce frequency tables of the disease and examine the related descriptive statistics. Mapping the data is also helpful.

Diseases will occur at different rates in different age, ethnic and sex groups. The calculation of expected numbers should take into account the possible effect of these possibly confounding factors on the occurrence of the disease. If the number of cases and population size are sufficient, confounding factors can often be accounted for using some form of standardisation, direct or indirect.



If the number of health events is too small to show meaningful rates, pooling across geographical areas or time may be possible. Other commonly used statistical approaches include chi-square tests of observed versus expected frequencies based on the Poisson distribution for low-frequency data and Poisson regression used for comparison of rates. Confidence intervals may be calculated for point estimates.

A space cluster can be evaluated by comparing the rate in the study area with that in adjacent census areas or changing the geographical scale at which the analysis is carried out – for example, local area, town, city, health district, territorial authority, region or New Zealand. If a time cluster is being assessed, the occurrence in that time period can be evaluated in the context of previous or subsequent periods. When comparisons are made, the comparison population must be carefully chosen to ensure it has similar demographic or exposure characteristics.

Analysis of cancer incidence data at a range of geographic scales gives information that can deal with public concerns, prevent expensive and unwarranted epidemiological studies driven by public and political pressure, and target appropriate cases for further investigation.

Statistical verification of alleged clustering may be very difficult. For rare events, the timeframe selected for determining the expected incidence will influence whether the cluster is found to be statistically significant. Short timeframes may be misleading. Defining geographic boundaries may have a similar effect.

Problems may arise from statistical techniques used to detect clusters (Rothman 1990; Wartenberg and Greenberg 1992). It is often difficult to distinguish between events of epidemiological and public health importance and those that result from chance. Some techniques may not be sensitive enough to detect true aggregations, while others may detect aggregations whose epidemiological and biological significance is difficult to interpret. In a review of results for four commonly used methods under two alternative hypotheses of environmental exposure, Wartenberg and Greenberg (1992) showed that false negative rates depended heavily on the method used and the nature of the exposure pattern that was sought. True clusters were detected only if enough cases had been observed and if the method most sensitive to the type of exposure pattern had been used.

It is desirable to obtain statistical advice. Many alleged clusters require only basic data analysis. For rare diseases in small areas, the alleged cluster may only be one or two cases and may disappear once case histories and diagnoses are verified, and recent migrants and other anomalies detected as each case is reviewed. Clusters may also disappear or reappear by changing the time, space or time–space boundaries, by over- or under-enumeration of the population at risk, or by choosing different sets of standard rates. Decisions are often implicit rather than explicit, as they depend on existing data (Elliott et al 1995). Results should be treated cautiously until more about the sensitivity and specificity of the methods used is known (Elliott 1995).

Assessment of chance is typically based on a retrospective cohort analysis of routine data, such as from a population-based cancer registry. Age-standardised rates in the exposed population are compared with the reference population (eg, population covered in the cancer registry). The result is an age-standardised incidence or mortality ratio. Confidence intervals can be calculated to assess statistical significance.

Coory and Jordan (2013) consider cluster investigation should focus on exposure assessment instead of the assessment of chance. Among their reasons for this view are that *p*-values and confidence intervals are impossible to interpret due to the large and unknown number of multiple comparisons (the Texas sharp-shooter fallacy), it is not possible to adjust for confounding factors other than age and sex, and migration occurs in and out of the area. They recommend the epidemiological aspect of cluster investigation becomes a case series. If a common exposure is suspected to be the cause of the case series, the causal relationship can be investigated with case-control or cohort studies (which involve confidence intervals).

Others consider chance should be assessed (Assunção 2013; Lawson 2013; Waller 2013). Methods are available to deal with the problem of multiple comparisons (eg, spatial scan statistic) (Assunção 2013; Waller 2013).

Despite the lack of consensus about the definition of a cluster, a variety of statistical techniques can detect one. Brief descriptions of some of the available statistical techniques are given in the CDC's 'Guidelines for investigating clusters of health events' and its addendum (CDC 1990; Kingsley et al 2007).

The United States National Cancer Institute has developed free spatial-scan statistical software, SaTScan.<sup>9</sup> It can be used to:

- evaluate reported space, time or time–space clusters and determine if they are statistically significant
- test whether a disease is randomly distributed over space, time or time–space
- carry out geographical disease surveillance
- test geographical areas of significantly high or low rates
- carry out surveillance for early detection of disease outbreaks (Kearney 2008).

Investigation of a cluster is difficult. Review results carefully in an effort to find patterns suggesting possible similarities linking the affected cases, which may form the basis of exposure hypotheses.

If there is no cause identified, then there is no objective evidence on which to obtain expert judgement about its importance. A *p*-value, even when adjusted for multiple comparisons,

<sup>9</sup> Available at [www.satscan.org](http://www.satscan.org) (accessed 6 September 2023).

does not give an objective measure of whether a cluster is due to chance. Ultimately, experts have to make a decision in the presence of uncertainty (Coory 2010).

## **Step 5: Make a decision – stop and report or investigate further?**

The decision on whether to stop and report on the investigation or continue to the next stage depends on a number of factors. Further investigation is usually not required if there is:

- no excess disease and no exposure, and therefore no biological plausibility
- no excess disease, a possible exposure, but no biological plausibility that the exposure could result in an excess
- excess disease, no identified exposure and no biological plausibility that the excess rate results from an environmental exposure.

If there is an excess of cases, consider the following questions.

- Is it of concern?
- Does it warrant further study?
- Is the exposure biologically plausible?
- Have cases increased suddenly in a recent period?
- Are cases more concentrated around suspected environmental hazards or in suspected occupational groups?
- Can the population at risk be defined?
- If cancer, is the type of cancer or age of onset unusual?
- If cancer, are there documented, prolonged exposures to known or suspected carcinogens at levels exceeding environmental limits?

In general, a 'yes' answer to most of these questions increases the need for further follow-up (Stage 4).

The threshold for further investigation varies among organisations and requires professional judgement. It involves finding an appropriate balance between scientific rationale and public concern (Benowitz 2008).

If the decision is made to end the investigation:

- write a report with a summary and conclusion
- obtain appropriate peer review

- communicate results to the public.

If further investigation is not warranted or feasible, what other actions could be carried out to address community concerns? For example, if a nuisance or undesirable level of exposure was identified, intervention that reduces exposure may be more important than proving a causal relationship.

## **Stage 4: Surveillance or epidemiological study**

If the cluster warrants further investigation, the two options are surveillance and conducting an epidemiological study.

### **Surveillance**

Surveillance is more appropriate than an epidemiological study where an excess number of cases is found in the cluster but it is of low statistical significance or the exposure has weak biological plausibility. A surveillance programme run over several years determines whether cases are increasing over time and what their geographical distribution is. Surveillance may also occur after an epidemiological study to monitor incidence.

If no registry or vital statistics data are available, a reporting system may have to be established to receive reports about the disease from the public or health professionals.

### **Epidemiological study**

If there is an excess of cases and they have a biologically plausible connection with some environmental exposure, further investigation of the cases and their environment is warranted.

Further investigation may involve a case-control, cohort or cross-sectional study and can range in duration from a few days to years of work and in budget from hundreds to thousands of dollars. Consultation with appropriate specialists and agencies is recommended. These include the Te Whatu Ora, university departments of public health, and the Institute of Environmental Science and Research.

# Legislation

A range of legislation may apply during a cluster investigation. Potentially relevant statutes and regulations<sup>10</sup> include the:

- Health Act 1956
- Resource Management Act 1991
- Health and Safety at Work Act 2015
- Cancer Registry Act 1993
- Privacy Act 2020 (and Health Information Privacy Code 2020)
- Hazardous Substances and New Organisms Act 1996
- Health and Disability Commissioner (Code of Health and Disability Services Consumers' Rights) Regulations 1996
- Accident Compensation Act 2001
- Local Government Act 2002.

The Privacy Act 2020 establishes the set of information privacy principles that govern the collection, holding, use and disclosure of information. The Health Information Privacy Code 2020<sup>11</sup> modifies the privacy principles to deal specifically with issues that arise in the health sector. It applies to health information related to identifiable individuals rather than anonymous or aggregated statistical information from which individuals cannot be identified.

The Code contains 13 rules that apply to health information and health agencies. Rules 1-5, 10 and 11, relating to collection, storage, use and disclosure, are the most relevant to cluster investigations.

The rules have a number of exceptions. Many are circumvented by the exception that, if the information is to be used for research or statistical purposes and it will not be published in an identifiable form, then the particular rule does not apply.

Exceptions to Rules 10 and 11 allow disclosure of information if the health agency holding information obtained in connection with one purpose believes on reasonable grounds that:

- its use or disclosure for any other purpose is necessary to prevent or lessen a serious threat to public health or safety, or

<sup>10</sup> Available at [www.legislation.govt.nz](http://www.legislation.govt.nz)

<sup>11</sup> Available at <https://www.privacy.org.nz/privacy-act-2020/codes-of-practice/hipc2020/>

- non-compliance with the rule is necessary to avoid prejudice to the maintenance of the law by any public sector agency.

The definition of 'serious threat' in the Privacy Act 2020 for the purposes of information privacy 10(1)(f) or 11(1)(f) means:

a threat that an agency reasonably believes to be a serious threat having regard to all of the following:

- (a) the likelihood of the threat being realised; and
- (b) the severity of the consequences if the threat is realised; and
- (c) the time at which the threat may be realised.

Disclosure in these circumstances must be made to a person who can do something about the threat and must only be to the extent necessary to prevent or lessen the threat.

When collecting information, it may not be practicable to contact the individual for consent or to get the information directly from them. The agency or informant providing the information must believe on reasonable grounds that one of the exceptions to Rule 10 or 11 applies or that a designated officer is collecting it according to statutory functions.

The Privacy Act 2020 and the Health Information Privacy Code 2020 are subject to other countervailing legislation. If designated officers are acting according to specific statutory powers, the Act and Code are overridden. Care has to be taken that the actions remain within the scope of those powers; otherwise, privacy legislation must be considered. If difficulties or concerns arise, consult with your organisation's privacy officer and/or legal advisor.

Designated officers conducting a cluster investigation should be familiar with their responsibilities under the Code of Health and Disability Services Consumers' Rights 1996<sup>12</sup> and should consider their investigation in light of the rights of (proposed) participants.

## Ethics

Cluster investigations for the protection of public health do not require health and disability ethics committee review. Public health staff conducting cluster investigations are, however, free to seek ethics committee advice about any special issues that might arise. If there is uncertainty about the need for ethics committee review, the relevant ethics committee should be consulted about the particular point(s) to which the uncertainty relates (National Ethics Advisory Committee 2019).

<sup>12</sup> Available at <https://www.hdc.org.nz/your-rights/about-the-code/code-of-health-and-disability-services-consumers-rights/> (accessed 6 September 2023).

The National Ethics Advisory Committee's *Standards for Health and Disability Research and Quality Improvement* (2019)<sup>13</sup> apply to both observational studies (which include cluster investigations and public health surveillance) and intervention studies. These Standards set out two sets of principles:

- Te Ara Tika principles – a set of Māori ethical principles that draw on a foundation of tikanga (Māori protocols and practices)
- Bioethics principles – which have been used in numerous sets of human research ethics guidelines.

The two sets of principles are the ethical sources of the more specific standards contained in the *Standards for Health and Disability Research and Quality Improvement*. When introducing the principles, the Committee notes that the Standards do not ethically or conceptually prioritise either of the two sets of principles. No assumption is made that they cover the same ground in all cases, but they do both involve knowledge discovery through respectful and rights-based engagement between researchers, participants, and communities to advance health and wellbeing. When used together, the two sets of principles address ethical positions of different societies, thereby strengthening ethical discourse in New Zealand.

A summary of the two sets of principles are shown in the Figure 5 below. For more information, refer to pages 34-37 of the Standards.

**Figure 5: Overview of Te Ara Tika and bioethics principles<sup>14</sup>**



<sup>13</sup> Available at: <https://neac.health.govt.nz/publications-and-resources/neac-publications/national-ethical-standards-for-health-and-disability-research-and-quality-improvement> (accessed 31 October 2023).

<sup>14</sup> Ibid – see pages 34-37.

# Risk communication

As already noted, it is often the public or media that alert the National Public Health Service to the occurrence of a cluster. The early stages of cluster investigation often depend on data provided by the public and media to build the file about the cluster.

Risk communication is a two-way process in which the National Public Health Service informs, and is informed by, the affected public. It involves an exchange of information about the nature, size and public health importance of a cluster and, if necessary, about the control of any health risk (Covello 1995).

The success of a cluster investigation does not depend on proving or finding the cause of a cluster, but rather on reaching a satisfactory outcome for all groups involved. Successful risk communication involves achieving mutual understanding and the resolution of any conflict between the public's expectation of a cluster investigation and scientific analysis within the limits of available data and knowledge (National Research Council 1989).

The Alberta Health Services (2011) guidelines suggest key communication milestones that are important to a successful risk communication strategy. These milestones are to:

- establish open communication with the informant and inform other agencies as necessary
- provide education (cluster investigation process, general disease information)
- identify the role of the informant in the investigation
- identify the investigation lead
- identify the working group
- communicate with the informant on an ongoing basis
- arrange for quality assurance – interagency evaluation and/or external peer review
- present results to informant.

Although the process is presented stepwise, some steps may occur concurrently and may be repeated (Alberta Health Services 2011).

An online toolkit for cancer clusters is available (National Public Health Information Coalition and CDC 2013). It includes a flowchart for handling related communication activities at each stage of cluster investigation, a communications plan template, how to use social media, and questions and answers. The toolkit can be adapted for local use and aspects are also relevant for non-cancer clusters.



# Perception of risk

Sandman (1991a) approaches risk as a combination of scientifically defined hazard<sup>15</sup> and public outrage. Therefore, the public and scientists usually have different estimates of risk (Slovik 1987). The risks that frighten and anger the public are not necessarily the same as those that kill people (Cohn 1989). People often feel threatened, powerless and outraged in response to a newly disclosed potential danger to their personal health, the public health and the environment (Rowan 1996).

A number of factors characteristic of non-communicable disease clusters increase public concern about risk to health. These factors include deaths and injuries that are grouped in time and space, considerable media attention, risks that are unfamiliar and uncontrollable, risks that involve an involuntary exposure and delayed effects, and situations where children are specifically at risk (Covello 1995).

The public and experts may use the same data, but come to different conclusions about possible environmental causation and levels of risk to health. A survey of the reactions of members of the public and epidemiologists to information about a hypothetical possible neighbourhood cancer cluster found that the public were not reassured by any neutral facts or most of the reassuring facts about the cluster and were more likely than experts to think that a cluster was present (Levy et al 2008).

The most powerful factors influencing the public perception of risk are the trust and credibility attributed to the source of risk information. This in turn is based on the source's perceived level of caring and empathy, competence and expertise, honesty and openness, and dedication and commitment (Covello 1995). In contrast, one of the least powerful factors is the scientific data about risk, although those data have usually been the main component of communications about risks and clusters.

## Communicating with the public

Communication with the public (and media) should begin early and as soon as possible after the cluster has been brought to a public health officer's attention. A public meeting to hear specific concerns and differing perspectives is often useful, but should be undertaken only after careful preparation and consideration of the relevant factors that the public are concerned about. This meeting gives an opportunity to explain what is known and what action is being taken and to give background information. The effectiveness of a meeting depends on the experience and credibility of the speakers (Thun and Sinks 2004). The public are more likely to agree on the findings of the investigation if they understand the types of information that are being sought and how different findings can be interpreted, for

<sup>15</sup>Sandman uses the term 'hazard' to represent scientifically defined risk (ie, the product of hazard and exposure). URL: [Covello](#) (accessed 6 September 2023).

example, by comparing community disease rates with corresponding data from New Zealand as a whole.

Investigators need to be able to recognise the source of any community suspicions of deliberate delay and cover-up, as well as the source of demands for unrealistic schedules and resource allocation (Rothenberg et al 1990). It is important to ascertain which factors that influence risk perception apply to the alleged cluster and to assess the social context. For example, how widespread is concern, how concerned are people, and are there other issues besides the cluster that people are concerned about (Drijver and Woudenberg 1999)?

In many instances, a careful explanation of the frequency with which health events occur or their randomness may allay many fears. Clusters, by their very nature, are unusual or unexpected events, but there is an identifiable probability that they will occur by chance. In addition, despite intense and robust investigation the cause of many non-communicable disease clusters has not been established. Failure to determine the cause does not reduce the impact of the cluster or its fear factor among the public.

Scientific reports, either oral or written, can either allay community fears about the cluster or create confusion, dissatisfaction and a call for continuing investigation. For example, a cluster investigation of spontaneous abortions found that risk perception and risk communication had an equivalent, if not, greater standing relative to the epidemiology and exposure assessment components of the study (McDiarmid et al 1994).

Communication between the National Public Health Service, public and media is always difficult when public outrage is high, risk is probably low and the science is uncertain. Sandman (1991b) offers epidemiologists eight guidelines for public communication.

- Tell the most affected people what you have found, and tell them first.
- Make sure people understand what you are telling them and what you think the implications are.
- Develop a mechanism to strengthen the credibility of your study and your findings.
- Acknowledge uncertainty promptly and thoroughly.
- Apply epidemiological expertise where it is called for rather than where it is unlikely to help.
- Show respect for public concerns, even when they are not based on science.
- Involve the affected people in the design, implementation and interpretation of the study.
- Decide that communication is part of your job and learn its basic principles.

These guidelines are similar to the following rules of risk communication from the United States Environmental Protection Agency (Covello and Allen 1988; Reckelhoff-Dangel and Petersen 2007).

- Accept and involve the public as a legitimate partner.
- Plan carefully and evaluate your efforts.
- Listen to the public's specific concerns.
- Be honest, frank and open.
- Coordinate and collaborate with other credible sources.
- Meet the needs of the media.
- Speak clearly and with compassion.

The public typically want certainty and simple, easy-to-read and comprehensible information from a credible source. Scientists may consider such messages as incomplete, inaccurate or even biased (Glanz and Yang 1996). For their part, scientists frequently deliver communications that are the converse of what the public is looking for: detailed technical statements that may be accurate and unbiased but are often complex and filled with jargon and uncertainty (Goldstein et al 1992).

While quantitative probabilities are intrinsic to cluster investigations, the public usually communicate in qualitative expressions. People often find it difficult to understand the technical meaning of terms such as 'unlikely', 'not statistically significant' or 'a probability of 0.05'. One study found that people were equally likely to interpret a '70 percent chance of rain' as 'rain 70 percent of the time', 'rain over 70 percent of the area' and '70 percent chance of some measurable rain' (National Research Council 1989).

Build trust and credibility by expressing:

- **empathy and caring**
- **competence and expertise**
- **honesty and openness**
- **commitment and dedication.**

(Covello 1995)

Public concern about clusters also gives an opportunity for education about certain diseases (eg, cancer) and their prevention and early detection. The following are some examples of potentially helpful messages about cancer.

- Cancer is common, affecting about one in three people in their lifetime in developed countries (The Lancet Oncology 2009).



# Relationships with the media

The media amplify public outrage but do not create it (Sandman 1991a). Journalists will usually want a story to have a visual component and to contain blame, politics, controversy and strongly emotive content. The media need to simplify complex and technical explanations and tend to avoid scientific qualifications or subtle distinctions. Newspaper coverage of clusters in the United States rarely mentions multiple adverse health outcomes, confounding factors, relative risk and data reliability. Scientists, on the other hand, usually believe accurate information about risks, generally couched in probabilities and uncertainties, should be the most important feature of a good news story and that more epidemiological and risk information should be included (Greenberg and Wartenberg 1990).

The media's role may change with the size of the community it serves. Newspapers in large metropolitan areas of the United States were more likely than newspapers in smaller communities to link contamination from local agents to threats to human health in the community and frame the story as problems. However, small community newspapers were more likely to frame local contamination in the context of solutions to the problem and to link contamination to health risks if the contamination was in a distant community (Griffen et al 1995).

Source credibility is also a major factor in developing a successful relationship with the media. If possible, get the support of environmental groups early in the investigation. For example, McCallum et al (1991) found that environmental groups were ranked consistently highly for perceived expertise and credibility. In contrast, officials in the chemical industry, who were the most knowledgeable about the risks of chemicals in the community, were the least trusted by the public.

Oral or written reports of cluster investigations need to be presented in a manner that will not lead the public or media to confuse or distort the message. The messages should be clear and unambiguous, use plain language, avoid jargon and reflect the perspective, technical capacity and concerns of the public. The key points need to be constantly and consistently stressed. Another approach is to provide background information on the health effect involved in the cluster as well as information on the scientific method before presenting the report of the study results (Curbow et al 1994).

The investigator should be straightforward and honest about fact, speculation and what is known, remain cooperative and responsive, and be prepared to provide any additional information rapidly. It is imperative that the investigator gains the trust and confidence of the public (and media) at the outset and does not promise what the National Public Health Service cannot deliver.

# Conclusion

The goal of risk communication is to establish trust and credibility. It is not a one-way process of informing the public of the scientific aspects of a cluster.

Responding to community concerns aroused by a cluster is integral to a cluster investigation. Early, consistent, honest and open communication can help to ensure that these concerns are addressed during the investigation and that inconclusive results do not come as an unwelcome surprise. It is important that affected communities know from the outset what a cluster investigation can and cannot achieve. Communicating information about the investigation's progress and outcome is vital.

The most important factor for successful risk communication is for the investigator to be respected as a credible and trustworthy source of information.

# References

- Abrams B, Anderson H, Blackmore C, et al. 2013. Investigating suspected cancer clusters and responding to community concerns: guidelines from CDC and the Council of State and Territorial Epidemiologists. *Morbidity and Mortality Weekly Report* 62(RR08): 1–14. URL: [www.cdc.gov/mmwr/preview/mmwrhtml/rr6208a1.htm](http://www.cdc.gov/mmwr/preview/mmwrhtml/rr6208a1.htm) (accessed 6 September 2023).
- Alberta Health Services. 2011. *Guidelines for the Investigation of Clusters of Non-communicable Health Events*. Edmonton: Government of Alberta.
- Alchin J. 1994. Investigation into birth defects in children of mothers working with pesticides for the Christchurch City Council. Unpublished report to the Christchurch City Council.
- Armstrong B, Aitken J, Sim M, et al. 2007. *Breast Cancer at the ABC Toowong Queensland*. Final report of the Independent Review and Scientific Investigation Panel. URL: <https://about.abc.net.au/wp-content/uploads/2013/04/BreastCancerABCToowongQLDFinalReportJune2007.pdf> (accessed 6 September 2023).
- Assunção R. 2013. Statistical assessment of cancer cluster evidence – in search of a middle ground. *International Journal of Epidemiology* 42: 453–5.
- ATSDR. 2002. *Case Studies in Environmental Medicine. Disease Clusters: An overview*. Atlanta: Agency for Toxic Substances and Disease Registry. URL: [www.atsdr.cdc.gov/hec/csem/cluster/docs/clusters.pdf](http://www.atsdr.cdc.gov/hec/csem/cluster/docs/clusters.pdf) (accessed 6 September 2023).
- Bandaranayake D, Read D, Salmond C. 1993. Health consequences of a chemical fire. *International Journal of Environmental Health Research* 3: 104–14.
- Bates MN, Lane L. 1995. Testicular cancer in fire fighters: a cluster investigation. *New Zealand Medical Journal* 108: 334–7.
- Benowitz S. 2008. Busting cancer clusters: realities often differ from perceptions. *Journal of the National Cancer Institute* 100: 614–15.
- Borman B, Read D. 1995. *The facts and fiction of a media cluster*. Presented to Annual Scientific Meeting, International Clearinghouse for Birth Defects Monitoring Systems, Helsinki, 1994. Abstract: *Teratology* 25: 122–3.
- California Department of Health Sciences. 1989. *Investigating Non-infectious Disease Clusters*. Berkeley: California Department of Health Sciences.
- CDC. 1989. Eosinophilia-myalgia syndrome – New Mexico. *Morbidity and Mortality Weekly Report* 38: 765–7.
- CDC. 1990. Guidelines for investigating clusters of health events. *Morbidity and Mortality Weekly Report* 39(RR-11): 1–16.

- Cohn V. 1989. Reporters as gatekeepers. In: M Moore (ed) *Health Risks and the Press*. Washington: The Media Institute.
- Coory MD. 2010. The ABC breast cancer cluster: the bad news about a good outcome. Letter to the editor. *Medical Journal of Australia* 193: 620–1.
- Coory MD, Jordan S. 2013. Assessment of chance should be removed from protocols for investigating cancer clusters. *International Journal of Epidemiology* 42: 440–7.
- Covello VT. 1995. Risk perception and communication. *Canadian Journal of Public Health* 86: 78–82.
- Covello VT, Allen FW. 1988. *Seven Cardinal Rules of Risk Communication*. Washington DC: US Environmental Protection Agency.
- Covello VT, Sandman PM, Slovic P. 1988. Part III Guidelines for providing and explaining risk comparisons. In: *Risk Communication, Risk Statistics, and Risk Comparisons: A manual for plant managers*. Washington DC: Chemical Manufacturers Association. URL: [www.psandman.com/articles/cma-3.htm](http://www.psandman.com/articles/cma-3.htm) (accessed 6 September 2023).
- Cox B, Skegg K. 1993. Contagious suicide in prisons and police cells. *Journal of Epidemiology and Community Health* 47: 69–72.
- Creech JL, Johnson MN. 1974. Angiosarcoma of the liver in the manufacture of polyvinyl chloride. *Journal of Occupational Medicine* 16: 150–1.
- Curbow B, McDiarmid MA, Breyse P, et al. 1994. Investigation of a spontaneous abortion cluster: development of a risk communication plan. *American Journal of Industrial Medicine* 26: 265–75.
- Department of Health. 1977. 2,4,5-T, spina bifida, and after. *New Zealand Medical Journal* 86: 99–100.
- Drijver M, Woudenberg F. 1999. Cluster management and the role of concerned communities and the media. *European Journal of Epidemiology* 15: 863–9.
- Driscoll T, Foster G, Driscoll F. 2008. *Investigation of a Reported Cluster of Cancer Cases at the National Gallery of Australia. Report to the National Gallery*. ELMATOM Pty Ltd. URL: <http://nqa.gov.au/Aboutus/OHS/Health/DriscollFinal.pdf> (accessed 12 May 2014).
- Elliott P. 1995. Investigation of disease risks in small areas. *Occupational and Environmental Medicine* 52: 785–9.
- Elliott P, Martuzzi M, Shaddick G. 1995. Spatial statistical methods in environmental epidemiology: a critique. *Statistical Methods in Medical Research* 4: 137–59.
- Elliott P, Wartenberg D. 2004. Spatial epidemiology: current approaches and future challenges. *Environmental Health Perspectives* 112: 998–1006.
- Elwood JM. 1990. Clusters, cataracts, and concerned citizens. *New Zealand Medical Journal* 103: 275.



- EUROCAT. 2007. *Statistical Monitoring Protocol 2007*. URL: <https://eu-rd-platform.jrc.ec.europa.eu/sites/default/files/eurocat-pub-docs/EUROCAT-Statistical-Monitoring-Protocol-2007.pdf> (accessed 6 September 2023).
- Fiore BJ, Hanrahan LP, Anderson HA. 1990. State health department response to disease cluster reports: a protocol for investigation. *American Journal of Epidemiology* 132: S14–22.
- Fleming LE, Ducatman AM, Shalat SL. 1991. Disease clusters: a central and ongoing role in occupational health. *Journal of Occupational Medicine* 33: 818–25.
- Fleming LE, Ducatman AM, Shalat SL. 1992. Disease clusters in occupational medicine: a protocol for their investigation in the workplace. *American Journal of Industrial Medicine* 22: 33–47.
- Gavin AT, Catney D. 2006. Addressing a community's cancer cluster concerns. *Ulster Medical Journal* 75(3): 195–9.
- Glanz K, Yang H. 1996. Communicating about risk of infectious diseases. *The Journal of the American Medical Association* 275: 253–6.
- Goldstein BD, Demak M, Northridge M, et al. 1992. Risk to groundlings of death due to airplane accidents: a risk communication tool. *Risk Analysis* 12: 339–41.
- Gould MS, Petrie K, Kleinman MH, et al. 1994. Clustering of attempted suicide: New Zealand national data. *International Journal of Epidemiology* 23: 1185–9.
- Greenberg M, Wartenberg D. 1990. Understanding mass media coverage of disease clusters. *American Journal of Epidemiology* 132(suppl 1): S192–5.
- Griffen RJ, Dunwoody S, Gehrman C. 1995. The effects of community pluralism on press coverage of health risks from local environmental contamination. *Risk Analysis* 15: 449–58.
- Hales S, Sabel CE, Exeter DJ, et al. 2005. Clustering of childhood asthma hospital admissions in New Zealand, 1999–2004. Presented at Spatial Information Research Centre annual colloquium, Dunedin, 2005. URL: [www.business.otago.ac.nz/sirc/conferences/2005/03\\_hales.pdf](http://www.business.otago.ac.nz/sirc/conferences/2005/03_hales.pdf) (accessed 16 October 2012).
- Herbst AL, Ulfelder H, Poskanzer DC. 1971. Adenocarcinoma of the vagina: association of maternal stilbestrol therapy with tumor appearance in young women. *New England Journal of Medicine* 284: 878–81.
- Juzych NS, Resnick B, Streeter R, et al. 2007. Adequacy of state capacity to address non-communicable disease clusters in the era of environmental public health tracking. *American Journal of Public Health* 97: S163–9.
- Kearney G. 2008. A procedure for detecting childhood cancer clusters near hazardous waste sites in Florida. *Journal of Environmental Health* 70: 29–34.

- Kiddle E. 2005. *Cluster Investigation into Motor Neurone Disease Nelson*. Nelson: Nelson Marlborough District Health Board.
- Kingsley BS, Schmeichel KL, Rubin CH. 2007. An update on cancer cluster activities at the Centers for Disease Control and Prevention. *Environmental Health Perspectives* 115: 165–71.
- Larkin GL, Beautrais AL. 2012. *Geospatial Mapping of Suicide Clusters*. Auckland: Te Pou o Te Whakaaro Nui. URL: <https://www.tepou.co.nz/resources/geospatial-mapping-of-suicide-clusters> (accessed 6 September 2023).
- Last JM (ed). 1988. *A Dictionary of Epidemiology* (2nd edition). New York: Oxford University Press.
- Lawson AB. 2013. Assessment of chance should be central in investigation of cancer clusters. *International Journal of Epidemiology* 42: 448–9.
- Lenz W. 1962. Thalidomide and congenital abnormalities. *Lancet* 1: 45.
- Levy AG, Weinstein N, Kidney E, et al. 2008. Lay and expert interpretations of cancer cluster evidence. *Risk Analysis* 28: 1531–8.
- Lucas RM, McMichael AJ. 2005. Association or causation: evaluating links between ‘environment and disease’. *Bulletin of the World Health Organization* 83: 792–5. URL: <https://apps.who.int/iris/handle/10665/269505> (accessed 6 September 2023).
- McBride WG. 1961. Thalidomide and congenital abnormalities. *Lancet* 2: 1358.
- McCallum DB, Hammond SL, Covello VT. 1991. Communicating about environmental risks: how the public uses and perceives information sources. *Health Education Quarterly* 18: 349–61.
- McDiarmid MA, Breyse P, Lees PSJ, et al. 1994. Investigation of a spontaneous abortion cluster: lessons learned. *American Journal of Industrial Medicine* 25: 463–75.
- Montana Department of Public Health and Human Services. 2006. *Montana Non-infectious Disease Investigation Protocol*. URL: [www.dphhs.mt.gov/publichealth/cancer/documents/CancerClusterInvestigationProtocol.pdf](http://www.dphhs.mt.gov/publichealth/cancer/documents/CancerClusterInvestigationProtocol.pdf) (accessed 25 June 2012).
- National Ethics Advisory Committee. 2019. *National Ethical Standards for Health and Disability Research and Quality Improvement*. Wellington: Ministry of Health. URL: <https://neac.health.govt.nz/publications-and-resources/neac-publications/national-ethical-standards-for-health-and-disability-research-and-quality-improvement> (accessed 31 October 2023).
- National Health and Medical Research Council. 2012. *Statement on Cancer Clusters*. URL: <https://www.nhmrc.gov.au/about-us/publications/nhmrc-statement-cancer-clusters> (accessed 6 September 2023).

- National Public Health Information Coalition, Centers for Disease Control and Prevention. 2013. *Cancer Clusters: A toolkit for communicators*. URL: <https://www.nphic.systems/toolkits/cancer-cluster> (accessed 6 September 2023).
- National Research Council. 1989. *Improving Risk Communication*. Washington DC: National Academy Press.
- Neutra RR. 1990. Counterpoint from a cluster buster. *American Journal of Epidemiology* 132: 1–8.
- O’Carroll PW, Mercy JA. 1990. Responding to community-identified suicide clusters: statistical verification of the cluster is not the primary issue. *American Journal of Epidemiology* 132(suppl 1): S196–202.
- O’Connell R, Parkin L, Manning P, et al. 2005. A cluster of thyrotoxicosis associated with consumption of a soy milk product. *Australian and New Zealand Journal of Public Health* 29: 511–12.
- Orioli IM, Mastroiacovo P, López-Camelo JS, et al. 2009. Clusters of sirenomelia in South America. *Birth Defects Research Part A: Clinical and Molecular Teratology* 85(2):112–18.
- Queensland Health. 2012. *Queensland Health Guidelines: Assessment of clusters of non-communicable disease 2012*. Brisbane: Queensland Health. URL: [https://www.health.qld.gov.au/data/assets/pdf\\_file/0018/442602/cluster-assessment.pdf](https://www.health.qld.gov.au/data/assets/pdf_file/0018/442602/cluster-assessment.pdf) (accessed 6 September 2023).
- Reckelhoff-Dangel C, Petersen D. 2007. *Risk Communication in Action: The risk communication workbook*. Cincinnati: Office of Research and Development, US Environmental Protection Agency. URL: <http://nepis.epa.gov/Adobe/PDF/6000012U.pdf> (accessed 6 September 2023).
- Rothenberg RB, Steinberg KK, Thacker SB. 1990. The public health importance of clusters: a note from the Centers for Disease Control. *American Journal of Epidemiology* 132(suppl 1): S3–5.
- Rothman KJ. 1990. A sobering start for the cluster busters’ conference. *American Journal of Epidemiology* 132(suppl 1): S6–13.
- Rowan F. 1996. The high stakes of risk communication. *Preventive Medicine* 25: 26–9.
- Rubin CS, Holmes AK, Belson MG, et al. 2007. Investigating childhood leukemia in Churchill County, Nevada. *Environmental Health Perspectives* 115: 151–7.
- Sandman PM. 1991a. Risk = Hazard + Outrage: a formula for effective risk communication (video). American Industrial Hygiene Association.
- Sandman PM. 1991b. Emerging communication responsibilities of epidemiologists. *Journal of Clinical Epidemiology* 44(suppl 1): 41–50S.
- Sare WM, Forbes PI. 1972. Possible dysmorphogenic effects of an agricultural chemical: 2,4,5-T. *New Zealand Medical Journal* 75: 37–8.

- Sitas F, O'Connell DL, van Kemenade CH, et al. 2010. Breast cancer risk among female employees of the Australian Broadcasting Corporation in Australia. *Medical Journal of Australia* 192: 651–4.
- Slovik P. 1987. Perception of risk. *Science* 236: 280–5.
- Smith AH, Duggan HM, Wright C. 1994. Assessment of cancer clusters using limited cohort data with spreadsheets: application to a leukemia cluster among rubber workers. *American Journal of Industrial Medicine* 25: 813–23.
- Smith D, Neutra R. 1993. Approaches to disease cluster investigations in a state health department. *Statistics in Medicine* 12: 1757–62.
- Stewart BW. 2007. 'There will be no more!': the legacy of the Toowong breast cancer cluster. *Medical Journal of Australia* 187: 178–80.
- The Lancet Oncology. 2009. Cancer clusters: how can we improve understanding? *The Lancet Oncology* 10: 1129.
- Thun MJ, Sinks T. 2004. Understanding cancer clusters. *CA: A Cancer Journal for Clinicians* 54: 273–80.
- Trumbo CW. 2000. Public requests for cancer cluster investigations: a survey of state health departments. *American Journal of Public Health* 90: 1300–2.
- Waller LA. 2013. Regarding assessments of chance in investigations of 'cluster series'. *International Journal of Epidemiology* 42: 449–52.
- Wartenberg D, Greenberg M. 1992. Methodological problems in investigating disease clusters. *Science of the Total Environment* 127: 173–85.
- Whorton D, Krauss RM, Marshall S, et al. 1977. Infertility in male pesticide workers. *Lancet* 2: 1259.
- Williams LJ, Honein MA, Rasmussen SA. 2002. Methods for a public health response to birth defects clusters. *Teratology* 66: S50–8.

# Glossary

Carcinogen	A substance capable of causing cancer.
Confidence interval	A range of values for a variable that has a specified probability of including the true value of the variable.
Confounding factor	A factor other than the exposure being investigated, which influences the outcome and so distorts a measure of the effect of an exposure on risk.
Designated officer	A medical officer of health, health protection officer or other officer designated by the Director-General of Health under the Health Act 1956 or other legislation.
Endemic	The constant presence of a disease or infectious agent in a given population group or geographical area.
Epidemiology	The study of the distribution and determinants of health-related states or events in specified populations.
Hazard	A situation or event of potential harm to health.
Latency	The time period between exposure to a disease-causing agent and the appearance of the manifestations of the disease.
Person-years	The number of years that a person in a study population has been observed.
Poisson distribution	A distribution function used to describe the occurrence of rare events or to describe the sampling distribution of isolated counts in a time or space continuum.
Relative risk	The ratio of the risk of disease or death among the exposed to the risk of disease or death among the unexposed.
Risk	The probability of harmful consequences arising from a hazard.
Risk assessment	The characterisation of potential adverse effects of exposures to hazards.
Risk communication	An interactive process of exchange of information and opinions among individuals, groups and institutions.
Sensitivity	A measure of the probability of correctly diagnosing a case, or the probability that any given case will be identified by a test.
Specificity	A measure of the probability of correctly identifying a non-diseased person.
Standardisation	A technique used to minimise the effects of differences in age when comparing populations.
Surveillance	Data collection to detect events or identify trends to initiate public health action.
Teratogen	An agent that produces birth defects in an embryo or fetus.

# Appendix: Cluster report form

## SUSPECTED CLUSTER REPORT

Name of person completing form

Date

### Informant

First name

Surname

Address

Telephone

Mobile

Email

Background of informant:

Media (*specify*)

Family member of index case(s) (*specify relationship*)

Friend of index case(s)

Doctor

Other (*specify*)

Description of the problem

How did the informant come to believe there might be a problem?

Setting of the suspected cluster:

Neighbourhood (*specify*)

School (*specify*)

Workplace (*specify*)

Other (*specify*)

Is there a suspected exposure?

Yes (*specify*)

No

Some of the following information about the index case(s) may be available from the informant:

Index case  (number, eg, 1, 2)

First name

Surname

Current address  
(last if deceased)

How long has the person lived there?

Age

Date of birth

 /  / 

Sex

Male

Female

Ethnicity

NZ European

NZ Māori

Pacific (*specify*)

Asian (*specify*)

Other (*specify*)

Diagnosis

Basis of diagnosis

Date of diagnosis

 /  / 

Date of death

 /  / 

Place of death

Suspected environmental exposures:

Type of exposure

Address where exposure occurred if  
different from above address

Date exposure began

 /  / 

Date exposure ended

 /  / 

Details of changes in exposure  
(eg, when, extent, duration)

Smoking history (year started, duration, amount/day, tobacco type (eg, cigarettes))

Occupational history:

Type of industry

Job

Year job began

Year job ended

Present job				
Previous job				
Job before that				
Job before that				

Any other details from informant

--

Is the informant willing to assist in providing further information if necessary?

- Yes
- No