ARCHIVED NEWSPAPER REPORTS AS A COMPLEMENTARY SOURCE OF EPIDEMIOLOGICAL DATA FOR RESEARCH INTO CLIMATE CHANGE ADAPTATION

Alberta, Canada, 1960–2000

COLIN L. SOSKOLNE, JUSTINE D. A. KLAVER-KIBRIA, KAREN M. MCDONALD, DONALD W. SPADY, J. PETER ROTHE, KAREN SMOYER TOMIC, KAILA-LEA CLARKE AND GIAN S. JHANGRI

Abstract

The impact of weather extremes on population health can be studied using descriptive information about pre- and post-event circumstances. Yet, descriptive data are not typically recorded in administrative databases used in quantitative health research by epidemiologists. This chapter introduces, describes and validates a method for using newspaper reports to complement traditional epidemiological data sources for research on climate change impacts.

As per the innovative methods and systems-based approaches advanced by Anthony (Tony) J. McMichael in his research across a broad array of epidemiological enquiry, our method development focuses on four areas by combining qualitative and quantitative methods: first, selecting and extracting information regarding extreme weather-related events and extreme weather-related disasters, and linking them to the appropriate newspapers; second, creating a content analysis framework (CAF) and extracting factual health data (i.e. manifest content) and its implied meaning (i.e. latent content) from newspaper reports; third, verifying the factual health-related data found in newspaper reports; and, fourth, corroborating the number of deaths cited in mortality data.
Our research supports the use of newspaper reports related to extreme weather as a complementary source of contextual epidemiological data when assessing climate-related health impacts to support more traditional epidemiological research in health and social policy.

**Introduction**

Human health and well-being are linked closely to social, cultural, economic and physical environments, which in turn are shaped and sustained principally by the forces of nature and human action.

The Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2007) states that over recent decades the frequency and severity of extreme weather events have increased, reflecting a trend towards global heating (Hansen, 2009) and more systemic and extreme climate disruption (Lovelock, 2009; Rockström et al., 2009). These changes have profound implications for life-sustaining habitats.

Extreme weather can disturb lives, livelihoods and physical and societal infrastructure profoundly. Direct health effects of extreme weather include injury, illness, significant human stress and mortality. Indirect health impacts include damage to local health services and civic infrastructure, disruption of community and economic functioning, and displacement of people (McMichael et al., 2003; Patz et al., 2005; Confalonieri et al., 2007; Berry et al., 2008; Lemmen et al., 2008; Costello et al., 2009).

Epidemiologists could help lessen the harms described above by assessing how communities cope with weather disasters, their emergency services and infrastructure, and local public health and social services. Understanding the nature and consequences of the direct and indirect effects of climatic disruptions is essential for developing effective risk assessments and policies for action and adaptation (Ebi et al., 2006; Berry et al., 2008).

In this chapter, we describe an innovative method (and its validation) for obtaining direct and indirect health impacts by using newspaper reports of the community impacts of and the responses to local extreme weather disasters. Such innovation is characteristic of Tony McMichael’s approach to public health challenges.
Aims and Objectives

We designed this study to assess (i) if newspaper reports could provide useful information for planning responses to weather events; and (ii) if qualitative methods applied to these data could be an effective analytic tool (Brown, 2003; Scammell, 2010).

Newspaper reports describe how significant weather events affect a community, plus they describe how community infrastructure and services have been affected; such qualitative data are normally not incorporated into quantitative epidemiologic analysis employing administrative databases. We evaluated and used data derived from newspaper reports to assess the impact of historical extreme weather events and disasters on Alberta’s population from 1960 to 2000. Such data could assist epidemiologists and health planners to gain insight into community responses to disasters associated with extreme weather and thus improve community disaster policy and planning initiatives.

Throughout this chapter, disasters are defined by the Canadian Disaster Database (CDD), and have specific criteria for inclusion. Extreme weather events, on the other hand, are calculated statistically using Environment Canada’s daily meteorological records as being outside of a community’s normal range of weather intensity, but do not meet the criteria for a disaster (both described below in the section headed Methodological Development). Unless specified, both ‘extreme weather events’ and ‘disasters’ are called collectively ‘extreme weather’.

Why newspapers?

Newspapers provide situational information (e.g. loss of critical infrastructure such as power during a heatwave) that could exacerbate negative health outcomes of populations affected by extreme weather. Such information is commonly unavailable in administrative databases. Newspaper reports often describe the number of individuals affected, including health outcomes, socio-economic conditions, infrastructure damage and other relevant contextual details of importance to public health. They also describe interviews on how the event and outcomes were experienced by locals.

The peer-reviewed literature using newspaper reports investigates mainly specific health issues; e.g. the media’s role in risk perception or in public health education. While newspapers commonly report research findings on health issues (e.g. Bubela and Caulfield, 2004), they are rarely used as a data source for health research. Consequently, valuable information may be missed.
Three relevant peer-reviewed studies combined disaster incidence with newspaper resources. Carley et al. (1998) found that most incidents in Britain went unreported in the health literature. Palecki et al. (2001) documented that, aside from mortality, newspapers reported additional impacts from heatwaves in the USA that would be inaccessible from secondary data sources (e.g. power losses, water shortages). Most recently, Stephens et al. (2007) demonstrated that newspaper reports of deaths in New Orleans after Hurricane Katrina were consistent with state-level vital statistics data. Thus, newspapers may help planners to anticipate the potential health and social consequences of future disasters.

Reporters describe the facts surrounding an event, editors set the tone and headline writers capture the essence. However, even ‘objective’ news reports reflect societal contexts and priorities, and the only ‘peer review’ is editorial oversight (Lowrey et al., 2007; Mogil, 2008). Therefore, our analytical approach includes a secondary level of enquiry. As described in the content analysis framework (CAF) below, we identified manifest and latent knowledge dimensions found within our newspaper reports, permitting exploration of their complementary nature for epidemiological assessments.

While mainstream reports should reflect a balanced perspective, editorials focus on issues and are often subjective (Fiske and Hartley, 2003). We wanted to gather contextual information, not editorial opinion, thus editorials were ignored.

**Content analysis of newspaper reports**

Media analysis is a rich source of public health information and is more commonly being used for analysis (Wilson et al., 2005; Kennedy et al., 2006; Renzulli et al., 2006; Barnes et al., 2008). While the latter three studies examined newspaper reports, their analyses focused on a single event; we consider the value of newspaper reports over extended periods and multiple events.

Human adaptation to climate change requires a multidimensional community response plan to decrease harms associated with future extreme weather. Therefore, researching the effects of climate change on human health would benefit from a mixed methods approach, because many dimensions of human society are affected. While some health consequences are described quantitatively by epidemiology, we argue that qualitative newspaper data will assist human adaptation to climate change. The data types are complementary and, used together, they can enrich the policy debate.

A newspaper report provides two forms of qualitative information: ‘manifest content’ (factual) and ‘latent content’ (hidden, institutional, interest driven) (Merton, 1949). According to critical communication theory (Rothe, 1993),
extracting and analysing both types of content using media analysis should liberate the reader from institutional influence. Indeed, Foucault and Rivière (1982) consider manifest information as insufficient, and we should focus on what influences the manifest, i.e. the report’s ‘latent content’.

Consistent with critical communication theory and qualitative research, we constructed a framework to address both manifest and latent contents in newspaper reports. In so doing, we make certain assumptions:

- As scientists we pursue truth. The media do not portray truth, but an ideological perspective on reality.
- The media, as an institution, mediates information flow. Our task as analysts is to deconstruct mediation rules.
- Information processed into knowledge is powerful. The analyst’s role is to sensitise the reader regarding the underlying ideology behind a newspaper report.

The above assumptions serve as the platform used in our CAF described below.

**Methodological Development**

Our approach used both quantitative and qualitative methods. We created a database combining quantitative weather data and administrative mortality data sources together with qualitative newspaper reports on ‘extreme weather’.

We used the CDD (Public Safety Canada, 2007) to select disasters (drought, heatwaves and coldwaves, snowstorms, forest fires and floods/rainstorms), and used Environment Canada’s daily meteorological records to identify extreme weather events (i.e. heat, cold, rain and snow). Tornadoes and hailstorms were excluded because, at the time, their association with climate change was unclear. We then selected and extracted reports from the newspapers of the community affected.

Finally, we developed a CAF to organise, code and assess health-related information reported in the selected newspapers. The accuracy of the reported health information was verified through various secondary data sources and the number of deaths cited in newsprint was compared to mortality data.
Disasters

A ‘disaster’ (Public Safety Canada, 2007) includes events meeting at least one of the following criteria:

- 10 or more people killed
- 100 or more people affected, injured, evacuated or homeless
- an appeal for national/international assistance
- historical significance
- significant damage or interruption of normal processes, such that the community affected cannot recover independently.

From the CDD, we identified 70 disasters across Alberta between 1960 and 2000; 19 drought and 51 non-drought disasters.

The CDD states the location of an extreme event. The newspaper used relating to each event was determined using four approaches:

- If the CDD identified an exact location, then the local community newspaper (usually published weekly) and the nearest daily newspaper were selected for investigation.
- If several communities were in the affected area, we investigated one or two randomly selected community newspapers and the nearest daily.
- If the CDD provided only a region, e.g. the Pembina River basin, then the ‘Rand McNally Alberta Provincial Map’ (2000) was used to determine communities in the area; from these, one or two local newspapers were selected randomly along with the nearest daily.
- Several extreme events, particularly drought, had prairie- or Canada-wide effects. For these events, all community newspapers associated with Alberta meteorological weather stations were identified. The province was then divided into three regions: north (seven communities), central (23) and south (25). One community was chosen randomly from each region. Since most community newspapers were weeklies, we selected the Edmonton Journal to provide a daily perspective reflecting provincial concerns. Edmonton was the most central major city and the capital of Alberta and was deemed most likely to cover provincial news.

The 40-year period (1960–2000) searched in newspapers for disasters adhered to the following search criteria:

- For dailies, we searched the duration of the event and two weeks beyond its end date.
- For weeklies, we searched for at least four weeks beyond the event end date.
• If the event was associated with only a weekly, we searched the weekly paper and the nearest city with a daily newspaper according to the rules outlined above.

• For droughts, the periods searched were modified so as to maintain a manageable scope and to address the fact that droughts had long-term effects and broad geographical ranges. Only months during which droughts had their most severe effects were searched, i.e. May–September, inclusive.

Of the original 70 disasters, 27 were unsearchable. Table 21.1 depicts the number of types of events identified in the CDD that were investigated originally and also the final number of disasters (by type) studied. For the remaining 43 disasters, 801 newspaper reports were analysed (Table 21.2).

Table 21.1 Number of each disaster type in the Canadian Disaster Database from 1960 through 2000 compared to the number of disasters analysed in this study.

<table>
<thead>
<tr>
<th>Disaster type</th>
<th>Number of disasters in the CDD</th>
<th>Number of disasters with newspaper reports that were investigated</th>
<th>Reason for exclusion (n = number excluded)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Floods</td>
<td>30</td>
<td>15</td>
<td>Overlapping disasters, vague dates and/or location (n = 15)</td>
</tr>
<tr>
<td>Cold</td>
<td>7</td>
<td>6</td>
<td>Vague dates and/or location (n = 1)</td>
</tr>
<tr>
<td>Snow</td>
<td>5</td>
<td>3</td>
<td>Vague dates and/or location (n = 2)</td>
</tr>
<tr>
<td>Fire</td>
<td>5</td>
<td>2</td>
<td>Vague dates and/or location, on a provincial border (n = 3)</td>
</tr>
<tr>
<td>Heat</td>
<td>1</td>
<td>0</td>
<td>Overlapping disasters (drought and flood) (n = 1)</td>
</tr>
<tr>
<td>Storms</td>
<td>3</td>
<td>2(^a)</td>
<td>Possible mistake in date, no reports found, vague dates and/or location (n = 3)</td>
</tr>
<tr>
<td>Drought</td>
<td>19</td>
<td>17</td>
<td>Possible mistake in database, prairie-wide event, larger problems in Saskatchewan (n = 2)</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>45(^a)</td>
<td>(n = 27)</td>
</tr>
</tbody>
</table>

\(^a\) Storms were later eliminated due to low numbers, lowering the total to 43.

Source: Authors’ work.
Table 21.2 Ratio of number of archived newspaper reports to weather-related disasters, Alberta, 1960–2000.

<table>
<thead>
<tr>
<th>Disaster type</th>
<th>Number of reports found in archives</th>
<th>Number of discrete events</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drought</td>
<td>137</td>
<td>17</td>
<td>8.1:1</td>
</tr>
<tr>
<td>Flooding</td>
<td>436</td>
<td>15</td>
<td>29.0:1</td>
</tr>
<tr>
<td>Cold</td>
<td>92</td>
<td>6</td>
<td>15.3:1</td>
</tr>
<tr>
<td>Snow</td>
<td>72</td>
<td>3</td>
<td>24.0:1</td>
</tr>
<tr>
<td>Fire</td>
<td>64</td>
<td>2</td>
<td>32.0:1</td>
</tr>
<tr>
<td>Total</td>
<td>801</td>
<td>43</td>
<td>18.6:1</td>
</tr>
</tbody>
</table>

Source: Authors’ work.

Extreme weather events

Extreme weather lies outside a locale’s normal range of weather intensity and is, therefore, uncommon. Although often destructive, this is not always the case (Francis and Hengeveld, 1998). We defined an ‘extreme weather event’ using upper and lower percentiles for respective types of local weather (heat, cold, snow and rain), as described below and in more detail in Soskolne et al. (2004).

Extreme weather events were extracted from the daily meteorological records provided by Environment Canada. For data to be entered into the Environment Canada Climate Archive (www.climate.weatheroffice.gc.ca/prods_servs/normals_documentation_e.html#appendixa), they must meet the World Meteorological Organization data collection and archival standards; i.e. over 84 per cent of data must be recorded for any acceptable climate station. There were 1211 stations in Alberta that had potentially usable weather data. A criterion of ‘30 years or more’ of observable data is required to establish climate norms for a local area; thus, only those stations with 30+ years of consistent daily weather records having at least four measured variables (specifically including temperature and precipitation) were considered. This reduced the number of ‘useful’ stations from 1211 to 102.

Communities selected for analysis had to be associated with a regular (daily or weekly) newspaper for the entire study period. Communities with more than one weather station had only one station included. Fifty-one communities met our final criteria. For extreme weather events, our analysis was limited to newspaper searches that extended to three days beyond each extreme weather event for dailies and to two weeks beyond the event for weeklies.
Extraction and assessment of health data

We used the CAF to determine the categories under which media reports would be analysed. The CAF reflected media analysis principles inspired by phenomenology and critical theory (Rothe, 2000). It used reasoning that depended on both deduction (categories were predefined or logical) and induction (categories were validated through a pilot study analysis of reports).

Qualitative researchers tend to look only at manifest meaning. In our framework, we considered both manifest and latent meaning, allowing us to establish the extent of any power relationships. Our CAF let us determine how newspaper reports demonstrated power relationships by examining the practices, techniques and procedures used in the reporting process. In doing so, we recorded how the reports were shaped, using the textual processes that underlaid the shaping of those reports (i.e. latent content), as well as the factual (i.e. manifest) content into one database.

The manifest content variables include descriptor and health-related variables. The descriptors included the newspaper studied and the location(s) and dates of the disaster. Health-related variables included advanced warning and preparedness, response, direct and indirect health outcomes, long-term solutions, mitigation or adaptation and additional information (for example, information about vulnerable segments of society).

Latent content variables included language use (the tone, metaphors, vernacular used in the report), forms of knowledge (political, economic, technical), explicitly stated beliefs and credibility of sources (who is quoted, why?; is the source credible?), case building, emphasis or ideology of the report (what is the key message?) and social relationships and social responsibility for the event (reports of consensus or conflict between major players).

We used the CAF to extract qualitatively and code numerically the information found in newspaper reports (Soskolne et al., 2004) and to show the ability of narrative reports in linking extreme weather events to health and well-being. Pilot studies were conducted to investigate the validity of the framework itself, to explore whether it could capture the required data and the reliability of newspaper search criteria.

The reliability of data abstraction was assessed by two raters, each independently examining a single sample of 150 newspaper reports, covering various events. They abstracted data on ten different aspects of each event, such as language use and advance warning, assigning codes to each. Because each newspaper report might have had several possible outcomes, the rater chose the dominant aspect first. We used Cohen’s Kappa to assess interrater reliability. Once determined, each report was evaluated by a single researcher using the CAF.
We verified factual health outcomes conveyed in newspaper reports by comparing five extreme events with secondary sources (see Appendix housed at www.colinsoskolne.com/documents/Supplemental-Published-Information.pdf). Further, we verified death counts to compare newspaper-reported numbers of deaths associated with three extreme weather events using newspaper reports against administrative mortality data from Statistics Canada as the benchmark. We found that both direct and indirect health outcomes were verifiable and that deaths could be readily verified for larger numbers, but not for smaller numbers.

This chapter focuses on the methods associated with using newspapers in epidemiological research. While some preliminary analysis has been completed (see Soskolne et al., 2004), this has yet to be formally analysed and published.

For the initial analysis of the manifest content variables, the data regarding health-related risks was consolidated into three categories: direct health effects, indirect health effects and contextual information. Direct effects include injury, death or illness. Indirect health effects include other determinants influencing people’s health negatively, such as energy and other service losses, infrastructure damage, water contamination and vulnerability of certain populations. Contextual information describes factors that contribute to health threats caused by extreme events and includes circumstances surrounding a particular event that can exacerbate its effects (e.g. a multi-year drought). Data from these three categories assist public health professionals as they plan and manage health risks associated with extreme events. This information also helped verify the accuracy of the data collected in the manifest content variables.

**Applications and Discussion**

In this chapter, we focus on the methods employed to determine the usefulness of newspaper reports for epidemiological enquiry. With an overall Kappa of 0.88, and Kappas ranging from 0.62 to 0.93 for individual aspects, the final CAF had a high degree of reliability.

An interesting observation from our study is the irregular ratio of the number of reports to disaster type in the following order of public interest: fire, floods, snow, cold and drought (Table 21.2). In Mogil (2008), ‘the newsworthy factor’ in reporting is recognised. This shows that, consistent with the observation by Mogil (2008), weather disasters need to be considered in context, such as their nature, degree of severity, economic costs and numbers of people affected.
Public policy

The type of data found in our analysis – descriptive, latent or health-related – could help formulate public policy regarding climate change adaptation. For example, past impacts could help assess the potential health and social consequences of future extreme weather, because the ways in which they were managed successfully could be incorporated into future response plans.

Individuals, organisations and institutions affected and responding to extreme weather include government departments, community committees, businesses, emergency responders, volunteer groups and community representatives. The contextual information revealed in newspaper reports, and not found in administrative data sources, may assist these groups to identify potential partners to help respond to future local extreme events. It may also help identify and coordinate their respective roles for future events. The database lends itself to future enquiry into the latent content dimensions of extreme weather. Future analysis of the latent content would be helpful.

Social planning

Our newspaper analysis demonstrated a broad range of health impacts and human experiences that could occur during extreme weather. These events, as noted in newspaper reports, can create intense psychological distress, health concerns, financial ruin, social upheaval and spiritual doubt.

Given the IPCC (2007) observation that extreme weather phenomena will become more common and more severe, preparing for effective adaptation measures is all the more critical. Thus, use of any data providing locally relevant contextual information to help develop adaptation measures, such as archived newspaper reports on past extreme weather experience, is recommended.

Although very resource-intensive, building the CAF was a valuable and successful exercise. The use of CAFs can provide a valuable tool in any epidemiological study where social context is relevant. Outside the scope of our analysis, a separate CAF could be constructed to examine how disasters and extreme weather events are editorialised in different media outlets. This approach would help determine social, cultural and political dimensions to a greater degree than from newspaper reports alone.

Limitations

Extracting disasters from the CDD and extreme events from Environment Canada meteorological records was straightforward in that they were both administrative data resources. However, extracting newspaper reports from various archives,
plus coding the reports, took far more resources than anticipated, mainly because of the time needed to search and code reports. Searching for newspaper reports in the archives was onerous; older records were of poorer quality, had smaller font sizes and were not indexed, and thus required manual searching. Incomplete library collections, short library hours and relocations added to the difficult task. Coding reports required a technician to scan them, convert the scanned reports to text and then edit the textual reports for accuracy. The researcher would read all reports, determine their relevance and then code each. It is likely that this intensive use of resources will decline with electronic media and searchable databases.

Although newspaper representations of the investigated events appeared to be accurate, they might be biased or incomplete. Newspapers have clear limitations, partly because other relevant events could be occurring but not reported. The tendency of newspapers to report other people’s opinions suggests that there could be unreported facts. What is reported may or may not assign all factors equal representation, possibly due to editorial policy. Major newspapers, like the Edmonton Journal, may more likely print national press accounts of disasters; thus reflecting the interests of large mainstream press agencies in their news coverage. Further, as editorial policy changes over time at the various newspapers, their reporting likely would require regular verification and cross-referencing with other sources.

Finally, while print news media provide a novel source for epidemiological research, the advent of electronic media may change the reliability and integrity of using future news reports. For retrospective events, archived newspaper reports have demonstrably reliable contextual sources. For future events, E-media could make such work more efficient, but (i) assessing the reliability of electronic news reports is essential, and caution, as with any data source, needs to be exercised when selecting sources; and (ii) electronic data storage and access may pose problems because such data sometimes are ephemeral.

In summary, we report two major methodological lessons learned regarding the use of newspapers as a data source for linking extreme weather events to health impacts. They are:

- Newspaper-based research takes longer than expected. There are usually more data than anticipated and older reports are more difficult to obtain and review.
- Formal training programmes should be created to guide archive researchers in meeting the challenges associated with newspaper research.
Advantages

The complementary aspects of (usually qualitative) newspaper-derived data will enrich data obtained from (usually quantitative) epidemiological data sources concerning climate change. Policymakers will find helpful the contextual aspects derived from newspaper reports. While newspapers could stand alone as a data source, the combination of both qualitative and quantitative data, by providing the context for traditional epidemiological data, will better inform policymaking.

We do not suggest that newspaper-based data eclipse traditionally obtained epidemiological data, but we do demonstrate that newspapers can be valuable complementary sources. The mixed methods approach provides complementary and context/content derivable insights, and may point to areas of potential research interest. Therefore, access to archived newspaper reports should be considered whenever insights are sought to address a complex health issue. This study is consistent with the holistic and systems-based approaches so characteristic of Tony McMichael’s many contributions.

Conclusions

We conclude that newspaper reports can be used as a source of complementary epidemiological data. The method described here may be applicable in other contexts. Because data collected from newspaper reports reflect local conditions, any lessons learned should be generalised with caution. We encourage the use of such data in much the same way that case reports act as leads to more rigorous, higher-level investigation.

Our CAF is reliable and serves as a model for others to conduct similar research. Using newspapers to achieve our objectives provides a perspective on the direct and indirect impacts of extreme weather in a prairie region. Using mortality data and secondary sources, we could verify the majority of health data found in newspaper reports, as also demonstrated by Stephens et al. (2007).

We conclude that newspaper reports are a reliable source of contextual data surrounding extreme events. Our study advances a method that is pioneering and, applied elsewhere, may well be valuable for informing public policymakers, emergency preparedness teams and health and infrastructure planners at the local, regional and national levels when developing adaptation strategies. Clearly, extreme weather events are multifactorial in both origin and effects, such that any analysis should reflect a systems-based approach.
Acknowledgements


References


This text is taken from *Health of People, Places And Planet: Reflections based on Tony McMichael’s four decades of contribution to epidemiological understanding*, edited by Colin D. Butler, Jane Dixon and Anthony G. Capon, published 2015 by ANU Press, The Australian National University, Canberra, Australia.