

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Health Impacts and Economic Costs of Residential Fires (RESFIRES study): Protocol for a population-based cohort study using linked administrative data
<b>AUTHORS</b>	Harvey, Lara A.; Ghassempour, Nargess; Whybro, Mark; Tannous, Kathy

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Kathryn Rodgers Silent Spring Institute United States
<b>REVIEW RETURNED</b>	03-May-2020

<b>GENERAL COMMENTS</b>	<p>This protocol provides a systematic way to collect information directly from 9 NSW, Australia data sources that have information on a fire incident and associated events, such as hospital visits. The protocol lists the five main research questions that this protocol hopes to answer. Because of the often incomplete nature of how fire incident data is reported, a systematic way to pull together relevant data from different sources is needed. Fire deaths are often preventable, and this protocol aims to identify ways to create prevention or intervention strategies to save loss of life from residential fires.</p> <p>One overarching question I have is, will the fire incidence dataset overcome the problem of lack of fire information discussed in the introduction? I am not familiar with the data in Australia, but in the United States, the lack of information provided in fire incident reports, or the lack of fire incident reports submitted altogether, require a different intervention, aimed at firefighters to complete data entry after an incident. Is there a way to identify data fields that have a high proportion of missing entries, to target for ways to improve data entry? One challenge to overcome is a high degree of missing information on a variable – it should be determined if the missing status is correlated with other fire characteristics, such as a small fire, or a rural location, or a service of volunteer firefighters with little time to complete data entry. Understanding the distribution of missing variables is important for determining how best to treat those variables in statistical analyses, that will ultimately answer the five important research questions in the manuscript.</p> <p>For research question 2, “fire related characteristics”, is there a way to collect information on smoke detector technology? Ionizing smoke detectors are designed to detect larger smoke particles from open flame fires, whereas photoelectric smoke detectors are designed to detect smoke from smoldering fires. Because smoldering fires are more deadly and produce toxic smoke potentially before flames are obvious, photoelectric or dual</p>
-------------------------	---

	<p>ionizing/photoelectric smoke detectors are preferred. A working smoke detector may still result in fire deaths if the fire started as a smoldering fire and the detector technology was ionizing. See “Investigating Smoke Alarm Effectiveness in Fatal Fires” by Fleming and Babrauskus for more information on this topic.</p> <p>Also related to “fire related characteristics”, there are other factors in addition to alcohol consumption that may make a person more vulnerable to a fire, such as a disability, or age. It would be helpful to include those variables as well. Will this section collect information on items that first ignited? In the United States, that has been a topic of research to prevent household items from igniting. There is increasing emphasis of non-toxic fire safety methods, given the harm to health caused by flame retardant chemicals. It would be helpful to publish this protocol along with the fields included in the fire incident reports filled out by the fire department.</p> <p>Another set of data that would be helpful to bring in are the smoking rates in the area of study. In the United States, cigarettes are still the leading cause of residential fires. Anti-smoking campaigns are not only beneficial to lung health, but also to protecting fire safety concerns. The ability to look at smoking related fires and smoking trends could help quantify additional benefits of smoking reductions (hopefully) in the population.</p>
--	--

<b>REVIEWER</b>	Ian Pike University of British Columbia, Canada
<b>REVIEW RETURNED</b>	22-May-2020

<b>GENERAL COMMENTS</b>	<p>The objective of this study is to quantify the incidence, health impacts, risk factors and economic costs of residential fires in New South Wales, Australia.</p> <p>The authors propose to link nine data sources to provide a comprehensive understanding of individual trajectories from fire event to first responder use (fire and ambulance services), emergency department presentations, hospital admissions, burn outpatient clinic use and death.</p> <p>The linked data will be used to describe the circumstances and characteristics of residential fires, provide a profile of fire-related injuries (including burns, smoke inhalation, other injuries sustained during fire) examine trends over time, and explore the relationship between fire circumstance, emergency and health services utilization, and health outcomes.</p> <p>Regression modelling, including multilevel modelling techniques, will be used to explore factors that impact on these relationships. The authors also propose to construct costing models.</p> <p>Overall, this reviewer judges the proposed methodology to be appropriate and relevant to addressing the research questions. While the language terminology is of a more general nature, this is perhaps expected prior to data linkage and the authors dealing directly with the resulting linked data set. The authors have expressed considerable thought as to how they will go about the analysis of what will be a huge linked data, providing appropriate detail on the linkage variables, and sufficient detail on the expected statistical analyses methods that will be utilized. This reviewer finds no red flags that would warrant rejection of this proposed protocol.</p> <p>Some recommendations for minor revisions are warranted, however:</p>
-------------------------	--

	<p>Firstly, the authors have not stated their understanding of the quality of the data sets, nor how they will assess data quality. The use of administrative data in other settings has resulted in the development data quality frameworks to ensure that the research analyses based on these data can be used to draw meaningful conclusions. The authors are encouraged to consider the use of an appropriate data quality assessment framework as part of the methodological approach to this study. While linking nine different data sets to achieve the objective of this study is entirely laudable, it is essential that the quality of that data allow for meaningful conclusions and recommendations.</p> <p>Secondly, the authors are encouraged to provide details on how they will deal with missing and unknown data which might also impact on the conclusions drawn from this study. Missing data is a problem in much research that utilizes even the most carefully collected administrative data. And, methods such as step-wise regression have been shown to be problematic when significant data are missing.</p> <p>Finally, the project is a huge undertaking, with various different foci. While I imagine it to be the case, the authors are encouraged to present the various analyses in multiple companion manuscripts rather than as one comprehensive study paper. In this way, a reader of this work will be given the opportunity to better understand how particular objectives were framed within a research method and how the study was undertaken. For example, the development of costing models will be a complex undertaking in and of itself, warranting a separate manuscript. Again, this may well have been considered by the authors, and is in fact their intended approach. If so, I encourage stating this as part of the proposed protocol.</p> <p>Thank you for the opportunity to review this study protocol. This is a large undertaking that has the potential to produce valuable knowledge related to the burden of fire related injury, and from that knowledge an understanding of what countermeasures might be appropriate to address the burden.</p>
--	--

<b>REVIEWER</b>	Charles Jennings John Jay College of Criminal Justice The City University of New York
<b>REVIEW RETURNED</b>	01-Jun-2020

<b>GENERAL COMMENTS</b>	The protocol is well-constructed, The research area is worthy and should make a valuable contribution to the literature. The methods are appropriate and the interdisciplinary approach is laudable and should assure applicability if finding both in research and practice.
-------------------------	--

### VERSION 1 – AUTHOR RESPONSE

**REVIEWER 1:**

1. Reviewer 1 expressed concern over the quality of information provided in fire incidence reports. “The lack of information provided in fire incident reports, or the lack of fire incident reports submitted altogether, require a different intervention, aimed at firefighters to complete data entry after an

incident.” ....“Is there a way to identify data fields that have a high proportion of missing entries, to target for ways to improve data entry?”

Response: We agree that the fire incident data may have a high proportion of missing values. One of the strengths of our research team is working in partnership with Fire and Rescue NSW. Co-investigator Whybro, for example, is an Assistant Commissioner at FRNSW, which provides the study team the opportunity to directly feedback data quality issues to FRNSW.

Assessing the quality of the data will be a paramount component of the cleaning and analysis process. On receipt of the data we will assess missingness, to determine whether data is missing completely at random (MCAR), missing at random (MAR) or missing not at random (NMAR) to ascertain whether principled methods, such as multiple imputation or inverse probability weighting are appropriate for use in our models.[1] Sensitivity analysis will be conducted around our results.

However, if a variable in the fire incident data proves to be too poor quality for meaningful analysis to be conducted, this will be documented. The manuscript has been amended to include our systematic approach to dealing with missing data. See Methods section, Data analysis plan, page 7.

[1] Perkins NJ, Cole SR, Harel O et al, Principled approaches to missing data in epidemiological studies. *American Journal of Epidemiology* 2018;187(3):568-575.

2. For research question 2, “fire related characteristics”, is there a way to collect information on smoke detector technology?

Response: The variable ‘Type of detector initiating Alarm’ does exist in the dataset, however as highlighted above, we are unsure whether it has been adequately coded. If it is of suitable data quality for use, then it is a factor we would consider in our analysis. Type of smoke detector has been added to the fire-related characteristics list. See Research question 1-2, page 8.

3. There are other factors in addition to alcohol consumption that may make a person more vulnerable to a fire, such as a disability, or age. It would be helpful to include those variables as well.

Response: Information on disability or age is not included in fire incidence data. We will explore person-level factors including age, comorbidity and disability, using the health service data (ambulance, emergency department, hospital and burn registry data) for the people who access health services as a result of residential fires.

4. It would be helpful to publish this protocol along with the fields included in the fire incident reports filled out by the fire department.

Response: There are 88 variables included in the fire incidents datasets which is too many to include in a Table in the protocol, further we are unsure of the completeness of data that will be provided. For these reasons we do not think it appropriate to provide a list of variables until after the data has been received and cleaned. We will include a supplementary table with variables in any publication that reports on the fire incident data.

5. Another set of data that would be helpful to bring in are the smoking rates in the area of study. In the United States, cigarettes are still the leading cause of residential fires.

Response: We agree that cigarettes are an important cause of residential fires. Smoking rates in NSW are slightly lower (11.2%) [1] than the US (13.7%). Unfortunately, we only have access to data on smoking rates from the NSW Population Health Survey that are ecologically based and summarised by health region, rather than post code or individual exposure level. For this reason, we have not explicitly mentioned smoking rates as a factor that we can evaluate in the protocol.

[1] NSW Government. Smoking in adults, NSW 2002-2019, HealthStats NSW. NSW Population Health Survey, NSW Ministry of Health. Available at:

[http://www.healthstats.nsw.gov.au/Indicator/beh\\_smo\\_age](http://www.healthstats.nsw.gov.au/Indicator/beh_smo_age)

#### REVIEWER 2:

1. The authors are encouraged to consider the use of an appropriate data quality assessment framework as part of the methodological approach to the study.

Response: Thank you for this suggestion, we will assess and report on the quality of the data we receive using the Australian Bureau of Statistics Data Quality Framework (ABS DQF).[1] The ABS DQF has been informed by the Statistics Canada Quality Assurance Framework (2002) and the

European Statistics Code of Practice (2005). The seven dimensions of quality that we will assess are: institutional environment, relevance, timeliness, accuracy, coherence, interpretability and accessibility. This has been added to manuscript, See Data analysis plan, page 7.

[1] Australian Bureau of Statistics. 1520.0 – ABS Data Quality Framework, 2009. Canberra, ABS. Available at:

<https://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1520.0Main%20Features2May%202009?opendocument&tabname=Summary&prodno=1520.0&issue=May%202009&num=&view=>

2. The authors are encouraged to provide details on how they will deal with missing and unknown data which might also impact on the conclusions drawn from this study.

Response: We agree, as per response to reviewer 1 point 1, we anticipate that there will be missing data particularly with respect to the fire data and have modified the manuscript to outline our systematic approach to dealing with missing data.

3. The authors are encouraged to present the various analyses in multiple companion manuscripts rather than as one comprehensive study paper.... I encourage stating this as part of the proposed protocol.

Response: It is our intention to publish a suite of manuscripts from the results rather than one large manuscript. This has been specified in the protocol, see Ethics and dissemination Section, page.11.

REVIEWER 3:

No revisions requested.

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Kathryn Rodgers Silent Spring Institute United States
<b>REVIEW RETURNED</b>	06-Jul-2020

<b>GENERAL COMMENTS</b>	The revised protocol addresses the questions previously raised around missing data, proposing to code missing entries as “missing completely at random (MCAR), missing at random (MAR) or missing not at random (NMAR)” to inform statistical analyses, including sensitively analyses. The protocol will also identify information on type of smoke detector, which is important because different smoke detector technology responds differently to fire types. Overall, the responses and revisions are well-reasoned and written to meet their main objectives of identifying targeted interventions to prevent fires and improve fire injury outcomes.
-------------------------	---