Supporting Information

Detailed methodology and supplementary results

This appendix was part of the submitted manuscript and has been peer reviewed. It is posted as supplied by the authors.

Disentangling health costs attributable to PM daily standard exceedances in Western Australia

To estimate health impacts and costs, we used a standard approach (1,2), with slight modifications, that is represented by the following figure:

Figure 1 Policy analysis approach used to estimate health impacts and costs. Adapted from (3)

The following sections detail specific values and methods used to estimate and quantify health impacts, apply a random forest model to classify smoke sources, estimate health costs, and perform a sensitivity analysis. Finally, we show how this method is practically applied with one simple theoretical example.
1. Estimation of health impacts

We estimated the number of cases using standard approaches for health impact assessments of exposure to air pollution (2,4). As detailed below we applied risk coefficients obtained from published epidemiological studies (5,6) to baseline incidence of the health outcomes of interest using a log-linear function described by the equation below.

\[ \text{Cases}_0 = IR_O \times \text{Pop} \times (e^{\beta_O \times \Delta C} - 1) \]

Where

- \( \text{Cases}_0 \): number of estimated cases for outcomes ‘o’
- \( IR_O \): base incidence rate for outcome ‘o’
- \( \text{Pop} \): estimated exposed population
- \( \beta_O \): health outcome risk coefficient for outcome ‘o’
- \( \Delta C \): change in PM\(_{2.5}\) concentration

We calculated the all-cause mortality incidence rate from state-wide death counts obtained from the Australian Bureau of Statistics (7). Circulatory and respiratory hospitalisation incidence rates were obtained from the online tables for Measure 1.04: Respiratory disease and Measure 1.05: Circulatory disease from the Aboriginal and Torres Strait islander Health Performance Framework (8). Asthma ED visit counts were obtained from the Emergency Department Care reports for 2014-15, 2015-16 and 2016-17 (9–11). Population data were obtained from the Australian Bureau of Statistics (12).

Table 1 Base incidence rates for mortality and hospitalisations (cases/100,000 persons-year)

<table>
<thead>
<tr>
<th>Cause</th>
<th>All-cause Mortality (*)</th>
<th>Hospitalisations - Respiratory</th>
<th>Hospitalisations - Circulatory</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>577.7</td>
<td>1,438.4</td>
<td>1,640.5</td>
</tr>
<tr>
<td>2003</td>
<td>576.06</td>
<td>1,443.2</td>
<td>1,666.1</td>
</tr>
<tr>
<td>2004</td>
<td>557.86</td>
<td>1,448.0</td>
<td>1,690.6</td>
</tr>
<tr>
<td>2005</td>
<td>563.84</td>
<td>1,453.6</td>
<td>1,712.4</td>
</tr>
<tr>
<td>2006</td>
<td>564.72</td>
<td>1,459.0</td>
<td>1,727.3</td>
</tr>
<tr>
<td>2007</td>
<td>585.43</td>
<td>1,464.9</td>
<td>1,735.3</td>
</tr>
<tr>
<td>2008</td>
<td>573.7</td>
<td>1,467.3</td>
<td>1,734.0</td>
</tr>
<tr>
<td>2009</td>
<td>559.4</td>
<td>1,470.1</td>
<td>1,736.4</td>
</tr>
<tr>
<td>2010</td>
<td>550.98</td>
<td>1,476.0</td>
<td>1,751.0</td>
</tr>
<tr>
<td>2011</td>
<td>534.88</td>
<td>1,478.4</td>
<td>1,763.5</td>
</tr>
<tr>
<td>2012</td>
<td>541.77</td>
<td>1,483.4</td>
<td>1,771.8</td>
</tr>
<tr>
<td>2013</td>
<td>535.35</td>
<td>1,488.5</td>
<td>1,781.1</td>
</tr>
<tr>
<td>2014</td>
<td>536.92</td>
<td>1,499.7</td>
<td>1,807.6</td>
</tr>
<tr>
<td>2015</td>
<td>562.36</td>
<td>1,511.8</td>
<td>1,838.9</td>
</tr>
<tr>
<td>2016</td>
<td>550.73</td>
<td>1,525.6</td>
<td>1,875.5</td>
</tr>
<tr>
<td>2017</td>
<td>550.73</td>
<td>1,541.2</td>
<td>1,916.2</td>
</tr>
</tbody>
</table>

*For 2017, we assumed same rate as 2016
We used the concentration response functions for the association between daily average PM$_{2.5}$ and health outcomes shown in Table 3: Multiple studies have documented a statistical increase in all-cause mortality and hospital admissions for respiratory and circulatory disorders. We use coefficients for these outcomes recommended by the World Health Organisation, derived from the meta-analysis from studies in comparable populations to Australia (5). This coefficient for the rise in mortality in response to air pollution is consistent with studies in many settings including wildfire smoke exposure in Australia (13), outdoor particulate air pollution in Europe (14), and North America (15), and in a study of tropical smoke haze in Malaysia (16).

There are a sufficient number of studies specific to landscape fire smoke and asthma to use a coefficient for asthma ED attendances and vegetation fire smoke exposure, based on a meta-analysis of international studies (6).

**Table 3 Selected dose-response functions**

<table>
<thead>
<tr>
<th>Effect</th>
<th>Cause</th>
<th>Age Group</th>
<th>Exposure</th>
<th>Beta (*)</th>
<th>Standard Error</th>
<th>RR per 10 (µg/m$^3$)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premature Mortality</td>
<td>All-cause</td>
<td>All</td>
<td>24h PM$_{2.5}$</td>
<td>0.001222</td>
<td>0.000393</td>
<td>1.0123 (1.0045-1.0201)</td>
<td>(5)</td>
</tr>
<tr>
<td>Hospital Admissions</td>
<td>Cardiovascular</td>
<td>All</td>
<td>24h PM$_{2.5}$</td>
<td>0.000906</td>
<td>0.000377</td>
<td>1.0091 (1.0017-1.0166)</td>
<td>(5)</td>
</tr>
<tr>
<td></td>
<td>Respiratory</td>
<td>All</td>
<td>24h PM$_{2.5}$</td>
<td>0.001882</td>
<td>0.001051</td>
<td>1.019 (0.9982-1.0402)</td>
<td>(5)</td>
</tr>
<tr>
<td>ED visits</td>
<td>Asthma</td>
<td>All</td>
<td>24h PM$_{2.5}$</td>
<td>0.00639</td>
<td>0.001344</td>
<td>1.066 (1.038-1.094)</td>
<td>(6)</td>
</tr>
</tbody>
</table>

(*) Beta estimated using the following equation: $\beta = \frac{\ln(RR)}{\Delta C} = \frac{\ln(RR)}{10}$

Exposure was estimated at spatial scale correspondent to Statistical Area 2 of the Australian Bureau of Statistics which have an average population ranging between 3,000 to 25,000 people (17), using an inverse distance weighting method (18) over daily PM$_{2.5}$ averages obtained from station data, whenever at least 1 monitoring station was within a 100 km range from the SA2 centroid. PM$_{10}$ and PM$_{2.5}$ hourly averages at station level were obtained from the Department of Water and Environment Regulation, Western Australia. We estimated daily PM averages using hourly data. Whenever PM$_{2.5}$ averages were not available for a specific station, we estimated PM$_{2.5}$ ($PM_{2.5} = PM_{10} \times PM_{2.5}/PM_{10_{ratio}}$), and prioritised the use of the $PM_{2.5}/PM_{10_{ratio}}$ using the following rule:

1) $PM_{2.5}/PM_{10_{ratio}}$ for each station and month
2) $PM_{2.5}/PM_{10_{ratio}}$ for each urban area and month
3) $PM_{2.5}/PM_{10_{ratio}}$ for each month

We considered an exceedance day, when either PM$_{10}$ or PM$_{2.5}$ daily averages were identified as exceedances in the NEPM reports.
2. Random Forest Model

To predict specific smoke haze cause (wildfire, prescribed burn or wood smoke) for the period 2008 – 2013 we applied a random forest algorithm using the ‘randomForest’ package (version 4.6-14) and the R software version 3.5.1 (19). Random forest is a method which uses a random sampling of a set of observations combined with machine learning algorithms to train a model and ‘grow’ classification ‘trees’ or clustered groupings of data, which are later used to predict classifications on new data (20). In our analysis, we used data with known specific smoke haze causes to train a model, and then applied this model to impute the most likely specific smoke haze cause on those cases where only a general ‘smoke haze’ cause was reported. We considered the following independent variables: station, year, month, day, day of week, urban area, daily PM$_{2.5}$, daily PM$_{10}$, daily minimum temperature, daily maximum temperature, daily average temperature, and average heating degree days (difference between a comfort temperature of 18 °C and daily average temperature). We obtained temperature data from the Bureau of Meteorology (21), for all stations that were active between 2002 and 2017.

We ran a sensitivity analysis on the number of trees (ntree) and the number of variables randomly sampled (mtry) to maximise model accuracy. Final model had a classification error rate of 11.11%.

3. Health Costs

Health costs were estimated using the following equation:

\[
Cost_{sO} = Cases_{sO} \times Unitary_{costsO}
\]

Where

- \(Costs_{sO}\): Total cost for outcome ‘o’ ($AUD)
- \(Unitary_{costsO}\): Unitary cost for 1 case of outcome ‘o’ ($AUD/case)

Mortality was valued using the value of a statistical life (VSL) as recommended by the Office of Best Practice and Regulation (22). VSL is commonly used to estimate the monetary benefits of reducing the risk of mortality, and is a measure of how much society values the reduction of one death. It does not take into account underlying health status or life expectancy. Hospitalisation costs were estimated using a cost of illness (COI) method, which includes the average cost of a hospitalisation obtained from the Independent Hospital Pricing Authority (23) National Cost Data Collection Cost Report, and the cost due to lost productivity represented by the lost income during the period of hospitalisation. We valued lost productivity as the average daily salary using the Average Weekly Earnings and Labour Workforce Statistics published by the Australian Bureau of Statistics (24,25). For ED visits, we used estimates from the Health Policy Analysis (26) Emergency Care Costing Report. We adjusted all costs to Australian Dollars of 2018 by applying inflation factors from the Reserve Bank of Australia (27).
### Table 4 Unitary health costs

<table>
<thead>
<tr>
<th>Effect</th>
<th>Cause</th>
<th>Value (AUD)</th>
<th>Year or Period of value</th>
<th>Inflation</th>
<th>Value 2018 (AUD)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premature Mortality</td>
<td>All-cause</td>
<td>4,200,000</td>
<td>2014</td>
<td>6.8%</td>
<td>4,486,062</td>
<td>(22)</td>
</tr>
<tr>
<td>Hospital Admissions</td>
<td>Cardiovascular</td>
<td>7,192.94</td>
<td>2015-2016</td>
<td>3.9%</td>
<td>7,473</td>
<td>(23)</td>
</tr>
<tr>
<td></td>
<td>Respiratory</td>
<td>7,279.99</td>
<td>2015-2016</td>
<td>3.9%</td>
<td>7,564</td>
<td>(23)</td>
</tr>
<tr>
<td>ED visits</td>
<td>Asthma</td>
<td>705</td>
<td>2015-2016</td>
<td>3.9%</td>
<td>732</td>
<td>(26)</td>
</tr>
</tbody>
</table>

#### 4. Sensitivity Analysis

Acknowledging potential limitations due to misclassification bias and error in the random forest classification model, we analysed impacts when excluding the period 2008 – 2013. As seen in Figure 2, the distribution of health costs is similar to the whole period analysis, with 52.7% attributable to prescribed burns, 37.5% to wildfires, and 9.8% to other reasons. Total costs are of AUD$ 110.8 (95% CI, 40 – 182.2) million, and an estimated 24 (95% CI, 9 – 40) premature deaths, 70 (95% CI, 40 – 103) asthma ED visits, and 160 (95% CI, 11 – 325) hospital admissions, for 139 days of exceedance. The years 2015, 2017 and 2006 are responsible for 26.4%, 24.8% and 10.2% of health costs during considered years.

![Figure 2 Estimated health outcomes and costs for the period 2002 – 2007 and 2014 - 2017: (A) Number of estimated cases per cause and particulate matter source, (B) Total costs (AUD 2018) by particulate matter source, and (C) Yearly costs (AUD 2018) by particulate matter source](image-url)
5. Practical example

We estimated the total number of cases and health costs for the theoretical impact of 10 (µg/m³) of PM$_{2.5}$ on a population of 5,000,000 people, for 1 day. We estimated a total of ~ 1 death, 6.4 CVD and RSP hospital admissions and 0.5 asthma ED visits, with a total cost of 4.2 $AUD million. When doing a sensitivity analysis for exposure, between 5 (µg/m³) and 15 (µg/m³), total cases range between 3.9 and 11.8, while costs range between 2 $AUD million and 6.3 $AUD million.

Figure 3 Estimated health outcomes and costs example
6. References


