Commentary on the Hazelwood mine fire and possible contribution to deaths

Professor Ian Gordon PhD, AStat, Director
Statistical Consulting Centre
The University of Melbourne

11 August 2015

Preliminaries

1. This report addresses instructions given to me by Ms Felicity Millner of Environmental Justice Australia. These instructions were contained in a letter dated 5 August 2015 (attached).

2. I was provided with the following documents to examine for the purposes of addressing these questions.

    1. Practice Direction No. 2, *Public Hearing for Terms of Reference 6.*
    2. Terms of Reference (refer to 6 and 7 only) dated 26 May 2015.
    3. Report prepared by VotV on the Births, Deaths and Marriages data, which includes the raw data received from BDM.
    4. Associate Professor Adrian Barnett's report.
    5. Department of Health analysis, *Reports of Deaths in the Latrobe Valley claimed to be related to the Hazelwood mine fire*, September 2014.
    6. Department of Health factsheet, *Reports of Deaths in the Latrobe Valley related to the Hazelwood mine fire*, 17 September 2014.
    7. Melbourne University, *Review of Birth Deaths & Marriages Victoria (BDMV) mortality data for the Latrobe Valley and the time of the Hazelwood coal mine fire in Morwell*, undated.
    8. Expert report of Professor Duncan Campbell.
    9. Email from Hazelwood Inquiry Board to VotV.
    10. Initial submission from VotV to Coroner dated 22 September 2014.
    11. Environmental Justice Australia submission to Coroner dated 29 October 2014.
    12. Extract from the 2014 Hazelwood Mine Fire Inquiry Report – Parts 4.1 to 4.3.

3. I am a Professor of Statistics and the Director of the Statistical Consulting Centre at The University of Melbourne. I have a PhD in Mathematical Statistics and am an Accredited Statistician of the Statistical Society of Australia Incorporated. I am a founding member of the Australasian Epidemiological Association. I have provided statistical consulting to several hundred clients from business, industry and government over the last 25 years. I am the author or co-author of about 70

papers in refereed journals. I have supervised or co-supervised four PhD students over the last ten years, and been Chief Investigator on ARC Discovery and Linkage Grants. I have appeared as an expert witness on statistical matters in numerous jurisdictions in Australia, including the Federal Court and the Australian Industrial Relations Commission. I was the President of the Victorian Branch of the Statistical Society of Australia in 2008-2009. I attach a brief CV to this report.

4. I assume that in general terms the readers of this report are familiar with the Hazelwood mine fire and the concerns raised about health. The essential details are described in my letter of instructions.

Data sources

5. The main data considered in the two reports I review here are deaths in four postcodes near the Hazelwood fire site, for the years 2009 to 2014 inclusive, and months January to June. This gives $4 \times 6 \times 6 = 144$ observations. These numbers of deaths are in a table in the document "V.O.T.V. Birth Deaths & Marriages (BDM) Death Statistics Latrobe Valley". The four postcodes differ considerably in population size and area.

6. The report by Louisa Flander and Dallas English states that these are the only data considered by them; they did not take local weather conditions into account, and age and sex distributions, and population movements, were not available.

7. These data also defined the outcome variable in the models reported by Adrian Barnett in his report. However, he supplemented his analysis with other data: population data for the La Trobe City Council, and temperature data at a monthly level; specifically, the maximum monthly temperature. His report does not indicate the sources for these extra data. The location he used for the temperature data may have been the La Trobe Valley weather station (station ID: 085280), for example. Further, he included other adjustments in his modelling, for season and trend, which I discuss below. These do not entail more data, but are designed to account for known or supposed time-related phenomena.

Flander and English report

8. The approach taken to analysis in the Flander and English report is to aggregate the deaths across the four postcodes. This is a reasonable strategy, assuming the absence of a clear ranking of exposure across the postcodes. (If exposure could be measured and differentially assigned to the four postcodes, a more refined analysis could be conducted by keeping the postcodes separate.)

9. Flander and English carried out a Poisson regression. Underlying Poisson variation is appropriate and standard for counts of cases of disease or death, since such data record events arise in a process occurring at a rate. In a Poisson

regression model, the rate is allowed to depend on potential explanatory variables.

10. It is relevant and desirable to specify the explanatory variables considered or used in any model reported. In the Flander and English report, this is not made clear. They mention several plausible explanatory variables that were not used, such as weather conditions and population size. They state explicitly that their models did not take external factors into account. At face value, it seems that a single explanatory variables was used, namely, 'month' as a categorical variable. I attempted to replicate their results (shown in their Table 1) using such a model but was unable to do so. In fact, if their model did just have 'month' as an explanatory variable, and no other terms, the predicted numbers of deaths in the 2014 months would simply be the averages of the respective numbers in the years 2009 to 2013, but the 'Predicted' numbers in Table 1 are not these figures. I investigated whether some other terms may have been used in the model, such as an overall trend with time, but was not able to find a plausible model that gave the 'Predicted' numbers in Table 1. I do not conclude that Flander and English have made an error in their Poisson regression analysis, only that it is insufficiently reported for the purposes of proper review.

11. I now wish to comment on the interpretation of the Poisson modelling carried out in the Flander and English report. Their Table 1 compares the actual numbers of deaths in 2014 months, with the numbers predicted on the basis of the years 2009 to 2013. They note that there were 37 more deaths in 2014 than predicted by the model, and that 'the additional deaths occurred in March and May'. In fact, for every month of 2014, the observed number of deaths was greater than the predicted number shown in their Table 1, to a varying degree. The lowest excess was a difference of +2, in January (before the fire).

12. The 'Lower bounds' and 'Upper bounds' of Table 1 are not described. I believe they have been derived as confidence intervals, probably 95% confidence intervals. If intervals are to be used to assess how unusual the observed numbers of deaths are, the appropriate intervals are not confidence intervals but prediction intervals.

13. A more direct method to assess the statistical significance of the observed numbers of deaths in Table 1 is to obtain P-values. A P-value is a way of representing statistical inferences; they are used in Table 2 of the Flander and English report. Effectively, we may ask: if the predicted number of deaths in February 2014 was 43.38, how surprising is an observed number of 50 deaths?

14. Calculations along these lines are shown in Table 1 below. The focus of both reports (Flander and English, and Barnett) is on the months of February and March, due to the dates of the fire. I consider it is reasonable to believe that any effect of the fire on mortality may have continued for some time after the fire was declared safe on 25 March 2014. It is not hard to envisage scenarios for which this is a logical possibility. A frail elderly person with chronic obstructive

pulmonary disease, for example, could have their respiratory system stressed by the air pollution from the fire in such a way that their death is accelerated, without it necessarily occurring during the period of the fire. For this reason, in Table 1 I consider a variety of time periods in 2014, starting with the two fire months separately, and then considering groupings of months, successively including more months. All the relevant predicted numbers are based on Table 1 in the Flander and English report. It is a feature of the Poisson distribution that the sum of statistically independent Poisson counts has itself a Poisson distribution, with rate equal to the sum of the individual rates. This property is used in Table 1 below.

Table 1: Comparisons of observed and predicted numbers of deaths in 2014, based on Table 1 in the Flander and English report.

Period	Predicted	Observed	Ratio	P-value
February 2014	43.38	50	1.15	0.175
March 2014	52.98	62	1.17	0.122
Feb – Mar 2014	96.36	112	1.16	0.064
Feb – Apr 2014	146.26	166	1.13	0.058
Feb – May 2014	199.24	228	1.14	0.024
Feb – Jun 2014	249.64	285	1.14	0.015

The last four of these P-values are small, and the last two are less than the conventional threshold of statistical significance, which is 0.05. The interpretation of the P-value is the probability of the observed number of deaths, or more, given that the predicted number of deaths governs the rate at which deaths are occurring. The smaller the P-value, the stronger the evidence that the 2014 death rates were abnormally high. Thus, on the basis of the numbers in Table 1 of the Flander and English report, there is quite strong and statistically significant evidence that the death rates from February to June 2014 were abnormally high.

15. Flander and English do not report P-values for the excess deaths in their Table 1, although they do report them for the alternative analysis they carried out, which was based on assuming an underlying Normal distribution for the variation in the counts of deaths. A more complete approach would have been to report them for both analyses.

16. An inadequacy of the analysis in (my) Table 1 is that it treats the predicted numbers as fixed, whereas they have actually been estimated from the 2009 to 2013 data. This can be corrected by fitting a Poisson regression model. Table 2 below shows the result of this analysis, in which the potentially different risk due to the fire is allowed to be in the same variety of periods as in Table 1; the only difference is that February 2014 and March 2014 are not separately considered.

Table 2: Rate ratios for Poisson regression models, using various periods of potentially different risk.

Period	Rate ratio	95% conf. int.	P-value
Feb – Mar 2014	1.20	(0.97, 1.47)	0.088
Feb – Apr 2014	1.16	(0.98, 1.38)	0.078
Feb – May 2014	1.18	(1.02, 1.36)	0.026
Feb – Jun 2014	1.17	(1.03, 1.34)	0.014

The rate ratios in Table 2 are similar to those in Table 1; as expected, the P-values are also similar but slightly larger, since the Table 2 analysis takes account of the sampling variation in the 2009 to 2013 data.

17. Flander and English provide an alternative analysis, assuming a Normal distribution for the underlying variation in the numbers of deaths. In my view the Poisson assumptions are to be preferred, although the Normal distribution may be a reasonable approximation. In the results of this analysis, shown in their Table 2, one of the analyses they describe compares January to June 2014 with the January to June periods of 2009 to 2013. Since the fire did not start until 9 February 2014, it is inappropriate to include January 2014 in any proxy measure of exposure to the fire.

18. In that analysis (Table 2, Flander and English) they report a predicted number of additional deaths per month of 9.2, for February to March 2014. This is a total predicted excess of 18.4. Note that this is of similar magnitude, but slightly larger than, the excess for the same period implicit in their Table 1, which is 15.6, obtained as the total difference between observed and expected in Table 1, for February and March 2014.

19. In overall terms, the report of Flander and English has found that there was an excess of at least 15 deaths in February and March 2014, compared with 2009 to 2013, from the area as a whole. This excess was not statistically significant at conventional levels of significance.

20. There are a number of ways in which the analysis could be refined.

Barnett report

21. The analysis in the Barnett report has some similarities with one of the two Flander and English approaches. Most notably, Barnett uses an underlying Poisson model for the variation in the death rates. I agree with this approach.

22. His analysis has a number of differences with the Flander and English model however. The major one is that he uses a Bayesian paradigm, which leads to a fundamentally different way of representing the results, although the two approaches can be sensibly reconciled. I comment more on this later.

23. Barnett has also adjusted for other phenomena that could help to explain some of the variation in the numbers of deaths. He has adjusted for population using La Trobe City Council figures, but whether or not this was accounted for 'had little impact on the results'. He allowed for an overall trend in death rates. He incorporated a seasonal term in his model, which is equivalent to adjusting for month, but in a way that assumes a smoothly varying effect over the course of the year. This was an appropriate way to adjust for time of the year. His analysis was at the postcode level and he fitted a random term for postcode, to accommodate the varying sizes of the postcode. Finally, he attempted to allow for temperature, in a simple way, by using the maximum monthly temperature 'from the Bureau of Meteorology'. These investigations of other potentially relevant phenomena are appropriate, in the attempt to estimate any fire effect.

24. The term allowed for the fire in Barnett's approach implicitly assumed that the potentially different risk of death from the fire arose in the two fire months (only). Two estimates are provided, with and without adjustment for temperature. These are a rate ratio of 1.14 without temperature in the model (Table 1, Barnett) and a rate ratio of 1.11 after adjustment for temperature (Table 2). The 2014 months of February and March were among the hottest in the series, which is why adjustment for temperature reduces the estimate.

25. I have analysed the data using the aggregated deaths, but in other respects the same terms as Barnett, and have obtained essentially the same results as he did, although I did not use a Bayesian approach. His 'probability that the death rate was not higher than the average during the fire' was found to be 0.11 without adjustment for temperature, and 0.20 after adjustment for temperature. These correspond to the P-values in the non-Bayesian approach.

26. He also estimated the excess numbers of deaths during the two fire months, and obtained 14.4 without adjustment for temperature; this is similar to the figure of 15.6 obtained by Flander and English. After adjustment for temperature, his estimate of the same quantity was 11.2.

27. The analyses in both reports are broadly consistent, even though they adopt different analytic paradigms and use different terms in their models. There was an estimated excess of deaths in the two fire months, and also in the subsequent three months of 2014. The quantification of this excess number of deaths (for February and March 2014) depends on the model used, and varies between about 11 and 18.

28. In all of these analyses, the excess is not markedly unusual according to strict conventions of statistical significance, in that the P-values are not smaller than 0.05.

29. As I have noted, if the possibility of a lingering effect on the risk of death is entertained, the excess risk does become statistically significant at such levels, for the period February to May 2014 and February to June 2014.

Department of Health documents

30. In my instructions I was asked to focus on the two documents from the Victorian Department of Health, the Report dated September 2014 and the Fact sheet dated 17 September 2014. It is a good feature of the Report that the data, at least summarised across months and postcodes, are presented clearly. However, in my opinion both of these documents lack an appropriate level of objectivity, as they focus on particular elements of the data and appear to be arguing persuasively towards a particular conclusion, namely, that the mine fire did not cause any excess deaths. For example, in the Report, there is a paragraph on the VOTV estimate of a 40% increase in deaths, and an attempted rebuttal. The next paragraph begins "Looking at the two months in which the fire occurred (February-March) there was a decrease of 19%". From the context, it would appear that this is referring to all postcodes, but in fact, as is clear from the Table, the sentence is referring to Morwell.

31. The Report notes that for Jan-June in 2014 the number of deaths in Morwell was 88, 'very similar to deaths in the years 2012 (89), 2010 (91) and 2009 (86).' This is selective reporting; the rest of the picture is that the 2014 figure of 88 was markedly higher than the other two years, 2011 (67) and 2013 (64).

32. The Fact sheet dated 17 September has content that overlaps with the Report and to that extent is subject to the same criticisms. Further, in the postcodes where there were excesses of deaths in the February to March period of 2014, there is either little discussion, or a comment prefaced by the word 'but'. In the case of Moe, there was an excess of 32%. The Fact sheet indicates that the Department is obtaining additional data to better understand this excess. It would be helpful to know what these data are.

Other issues

33. In the V.O.T.V. document it is noted that in the black Saturday fires of 2009, '11 people died in their homes in February'. It is not clear to me whether these 11 deaths are included in the data analysed, but if they are, there is at least a question whether they should be. Deaths from a special cause, particularly one at a particular time-point, do not reflect the natural variation in death rates which is of interest as a background comparison to the possible mine fire effect. I do not suggest that this is a clear-cut matter: deaths due to a bushfire could, from an alternative perspective, be seen as part of the elevated risk of high temperatures.

34. I note that the Government received advice from Monash University researchers that "no additional deaths would be expected even if the level of exposure to the measured level of air quality continued for six weeks". Six weeks was the approximate duration of the fire. The Monash University report was a substantial document which I am not formally reviewing here. It is my

understanding and belief that this assertion is based on an integrated exposure-response analysis of many studies. If it is the case that in these studies the exposure recorded accumulated over a long period of time, gradually, this is a different kind of exposure than that of the air pollution arising from the mine fire. A short, sharply elevated exposure, which clearly occurred in the Hazelwood fire, is a different matter. The famous London Smog of 1952 lasted for only five days but has been estimated to have contributed to an excess of several thousands of deaths. I am not suggesting that the Hazelwood fire's levels of pollution were similar to that London event, only that it is possible for a short air pollution exposure to have lethal effects. Further, it is widely accepted that the London event had an adverse effect on mortality in the months following.

35. The outcome data analysed are "all-cause mortality" counts. It would be worth considering refining this to cause specific mortality, as there would be some causes that could be ruled out as possibly due to the air pollution.

36. The Hazelwood Inquiry report noted that a 65% of Morwell residents received a relocation or respite payment (page 370). I am not sure what effect that might have on the data analysed, since I have not investigated the precise way in which deaths are attributed to postcode. More broadly, movement of people, associated with the fire, would be worth understanding better.

37. The data are analysed by month. In principle, a more refined analysis by day could be considered, for two reasons. Firstly, the fire did not start until 9th February, so deaths in February before that date should not be considered as plausible outcomes of it. Secondly, the Environment Protection Authority measurements, reported in the Hazelwood Inquiry report, offer the potential to examine the association between exposure and outcome in a more fine-grained way, if the deaths were available by actual date.

Conclusion

38. There is no doubt that air pollution can contribute to death; this has been comprehensively studied. So this situation is different from other cases where a cluster of cases of disease, or deaths have been noted, and concern raised about a possible cause. Here the possible cause is manifestly unambiguous. The question is, do the data demonstrate a strong enough association between the mine fire and mortality, to conclude that, in this case, the fire did actually contribute to deaths in some cases?

39. In reviewing the two reports that analysed the data, I conclude that they arrive at broadly similar conclusions, which is that there was an excess of deaths in association with the fire, of between 11 and 18 deaths, approximately, on the basis of comparison with the previous five years, in the area of interest. For February and March (the actual fire months) this excess is not statistically significant at conventional levels. This means that the data are consistent with general background variation, and no special effect of the fire. The data are also

consistent with the fire causing some excess deaths. The Barnett report, and my own analysis, suggest that the apparent fire effect is partly, but not entirely, due to hot temperatures, in that after adjustment for temperature the rate ratio for the fire effect is reduced slightly.

40. Based on my own analysis, in which the period of potentially different risk is assumed to extend beyond the actual time of the fire, (for example, to May 2014), the excess of deaths is statistically significant at conventional levels.

41. I have outlined the limitations I see in the data, and possible further lines of inquiry, in the body of the report, and especially in the "Other issues" section.