Author's response to reviews

Title: Vegetation fire smoke, indigenous status and cardio respiratory hospital admissions in Darwin, Australia, 1996-2005: a time-series study.

Authors:

Ivan C Hanigan (ivan.hanigan@anu.edu.au)
Fay H Johnston (fay.johnston@utas.edu.au)
Geoffrey G Morgan (Geoff.Morgan@ncahs.health.nsw.gov.au)

Version: 4 **Date:** 22 April 2008

Author's response to reviews: see over

RESPONSE TO REVIEWERS

Vegetation fire smoke, indigenous status and cardio respiratory hospital admissions in Darwin, Australia, 1996-2005: a time-series study.

Ivan Hanigan, Fay Johnston, Geoff Morgan

22.4.08

INTRODUCTION:

In response to the reviewer critiques, we have revised our manuscript. The comments were extremely valuable and we have made changes to address the concerns raised. Direct responses to individual comments are described below.

EDITOR

1. Add some numbers to the abstract to illustrate the tendencies observed.

We have added our estimates for the association of PM_{10} with total respiratory admissions (with no interaction term between indigenous status and PM) and the different associations estimated for the two population subgroups for respiratory infection admissions and total cardiovascular admissions on page 2 (line 20) and page 3 (lines 1-9).

2. The last sentence of the abstract can be left out.

This sentence (page 3, lines 12-14) was amended rather than deleted in response to reviewer 1 (minor essential revisions point 2) to acknowledge the limitation of our use of estimated rather than measured exposure data.

3. Please use sulfur instead of sulphur. Please add a section on Abbreviations.

Sulphur changed to sulfur on page 5, line 21. Abbreviations section added (page 23, line 6-7).

4. The title should include study design.

The title has been amended to "Vegetation fire smoke, indigenous status and cardio-respiratory hospital admissions in Darwin, Australia, 1996-2005: a time-series study".

5. Please indicate full institutional addresses for all authors.

Full addresses have been added.

6. On pg 11, use referencing for link.

We have now added this link (cited on page 13, line 21) to the reference list.

7. References 4 & 5, 7, 11, 13, 14, 20, 21, 27, 39 require modification (see placing of "In"), also in book chapters or articles in a book, include-Place of publication: Publisher; year: pgs). Remove issue numbers, "et al" is not appropriate, please indicate all authors.

The references have been amended to suit the instructions. We apologize that we did not find the instructions on how to cite reports (10-Koopmans, 13-Ellis, 15-Gras, 19-Australian Bureau of Statistics, 21-Trewin, 22-Environment Protection and Heritage Council, 25-Health Effects Institute, 31-Medina, 43-Samet) or declarations from conferences (11-Association for fire ecology congress) so these were formatted as books. Please advise if these should be formatted differently and we will happily oblige.

8. Figure legends should include a short title (max 15 words) and a detailed legend (up to 300 words).

These have been changed to conform with this requirement (see page 32)

9. In tables, ensure all cell lines display as black lines. Table titles should be a maximum of 15 words.

This issue has been addressed. Tables have changed and table 1 has been replaced by a different table.

10. Please re-upload figure 3, which we had difficulty accessing.

This image has been reformatted for this re-submission.

REVIEWER 1

Major Compulsory Revisions

1. Additional justification of the use of the predictive model to infer retrospective daily levels of PM_{10} is needed (including comment on the appropriateness of including humidity in the Poisson model of health effects also used to predict the PM_{10} level for that day).

We have expanded the methods section pages 8-11 to include more information about how the predictive model for PM_{10} was developed and validated. While these are

reported in detail in the original paper regarding the development of the model [1] we have now included more information in our paper. These extra details are also set out in detail in response to reviewer 2, point 1 (see below).

In summary the methods section now includes:

- Details of all available empirical data and how these were measured,
- Details from previous studies that have directly measured and compared PM levels at different sites within and close to Darwin demonstrating similar exposure over relatively wide regional areas,
- Details of which data were used to develop the predictive model,
- Specification of the predictive model,
- Details of measured and predicted daily data for 2005,
- Comparison of measured and predicted monthly averages for 1995,
- Comparison of predicted peaks in PM_{10} with satellite and written records of bushfires in 2000 and 2001.

The reviewer also expressed additional concern regarding the appropriateness of including weather variables in the Poisson model of health effects when they are also used to predict the PM_{10} level for that day. We feel that this is valid as the only term in both models is humidity; and its contribution to the PM estimate is small. Also it had no relationship with admissions in our models – we include it just to be consistent with previous studies.

2. Some comment on the use of discharge data, rather than emergency department visit records would be helpful.

We agree that for many conditions, and particularly asthma, many people present to the ED for treatment and are subsequently sent home without being admitted and that examination of ED data provides different and useful information. However we examine hospital admissions because these data were available for a ten year period while emergency department data are only available for a three year period. We have included a comment regarding this issue in the discussion on page 16, lines 20-22 and page 17, line 1.

3. Could the authors explain their use of a time-series methodology rather than following the case-crossover method.

Our previous case-crossover study had identified some associations between PM_{10} and hospital admissions in Darwin. However, particulate data were only available for three fire seasons and the precision of the estimated magnitude of association for all outcomes examined in that study was low. Here we attempt to address the statistical limitations of the previous study by using hospital admissions data over a ten year, rather than a three year period. We used the analytical approach of time-series modelling because, in direct comparisons with case-crossover analyses, this has been shown to provide slightly more statistical power resulting in smaller confidence intervals.[2, 3] Examination of the same outcomes using different exposure estimates and different analytical methods also provides the opportunity to examine the consistency of findings with two different approaches.

4. It would be helpful if the authors could put the importance of their study in context with results from other studies that have looked at wood smoke (see : Boman Scand J Work Environ Health. 2003 Aug;29(4):251-60). Specifically, in the introduction section of the paper, a statement as to what this study adds above and beyond other previously published studies would be helpful.

We have added text to highlight the paucity of evidence in this area and specifically referred to three relevant studies (including the one suggested) on page 4, lines 11-15. In addition we include a statement on page 5, lines 1-2 that highlights the contribution of our study because the city of Darwin enables the health effects of vegetation fire smoke to be assessed given the source of the PM pollution is due almost entirely to fire smoke.

Minor Essential Revisions

1. Abstract - methods: the assignment of air pollution on a daily level is critical to this study; the author should better describe how daily levels were assigned in the methods section of the abstract

We have revised the methods section in light of these comments. See detailed response to reviewer 2, point 1 below.

2. Abstract - results: some specific findings show be presented in the abstract to the reader can gauge the magnitude of effect and statistical significance of the study's key findings

Specific results are included in this section of the abstract on pages 2 and 3.

3. Abstract – conclusions: should acknowledge limitations of exposure data i.e., exposure to PM was not measured, but rather predicted from other variables with a modest level of correlation

We have included a sentence in the conclusions of the abstract on page 3 that acknowledge this study was limited by the use of estimated rather than measured exposure data.

Discretionary Revisions

1.1 I am unsure of the value of presenting age specific population counts and proportion of visits under the age of 15 when no age-specific analyses are performed.

This information was included to explain why terms for school holiday periods were included in some analyses but not in others. For instance some conditions have very few children were admitted, while for others, children constituted a major proportion of admissions, and for these we wished to account for the influence of school holidays on hospital admissions. Table 1 has been removed but the percentages of cases less than 15 years of age have been retained in Table 3.

1.2 The reviewer also asked if we could "describe/investigate whether children and the elderly are more susceptible to the effect of outdoor air pollution in this study population? Also on a related point, can such (age) difference in demographics contribute to differences found between the indigenous and non-indigenous populations?"

This is an important point raised by the reviewer. We did not investigate age related associations because of the extremely low numbers of daily admissions this would create for some diagnosis groups. This would disable our ability to investigate the differences between indigenous and non-indigenous subgroups. A key strength of our study location is that the population includes a relatively high indigenous population compared with other Australian cities, and that this is broadly indicative of low socio-economic status.[4] It is true that the two population subgroups have differing age

structures, with a greater proportion of people over sixty-five in the non-indigenous group, and a greater proportion of children less than 15 years in the indigenous group. [5] The former factor could result in an underestimate of the difference between the two population groups, while the latter could have contributed to the differences in respiratory infections observed between the two groups. As we have attempted to control for this age-structure effect by including a term for indigenous status which should capture this, we believe that the association of day-to-day PM_{10} variations is evident above and beyond this. Text has been added stating this on page 22, lines 4-10.

2. in tables, authors use the term PROPORTION in the titles, but present PERCENTAGES (Table 1, 3)

Titles have been changed.

3. table 4, is the influenza a rate per day or week? (in the text, I believe they mention it was calculated on a weekly basis)

This is correct. We used the weekly influenza rates (cases per 1000 consults) for each day of the week. More detail is now included in table 4 to describe this.

4. table 3 would be much more informative if the denominator (population) information for Indigenous and Non-indigenous groups were presented.

The Population in each group has now been included on the top row of this table.

5. Figure 2. The title is somewhat misleading. The line and dots, in themselves, I feel does not validate the use of the predictive model to do time series analyses for PM10.

We have revised figure 2 (now figure 1) to address these comments and similar concerns raised by Reviewer 2 (major revision 1). Figure 1 section A shows the modelled PM_{10} superimposed over actual PM_{10} for the period assessed using the validation dataset from 2005 (which was withheld from model development for the purpose of validation). This shows how well the day-to-day variations in PM_{10} are estimated, and the agreement between the minimum and maximum values. We include a footnote that the actual PM_{10} is include for comparisons purposes only, and that the study only used predicted PM_{10} value. Figure 2 section B shows the scatter plot of the predicted and observed values, (without the fitted line now as this was felt to be confusing).

REVIEWER 2

Major Compulsory Revisions

1. The methods section on exposures needs more detail. Because the exposures for this work are estimated, rather than monitor values, this is critically important.

We agree that our use of predicted data needed to be better described. In the methods section, pages 8-11 we have included text that explicitly describes the predictive model for deriving exposure measures for ambient PM_{10} from visibility data. We believe this is justified because of the limited availability of empirical air quality data and the fact that vegetation fire smoke is the main determinant of visibility during the dry seasons as rain and fog are rare events and there are no other important sources of air pollution. Darwin also provides the rare opportunity to investigate the effects of vegetation fire on health in the absence of other sources of particulate.

1.1. Input data: What type of input data were used and available, for all inputs, such as weather and PM10? (location of monitors, frequency of measurement (daily?) and timeframe of measurement).

This is now described on page 9 (6-23) and 10 (1-13) under the sub-heading: ***Data used for the development of the model.***

1.2. Rationale: If PM10 data are available, why use estimates from satellite data? I assume this is to fill in gaps for which no monitoring data are available; please specify the reason.

While we investigate the use of satellite data to supplement our PM data during the study period we did not use satellite data in the analysis. We also refer to an analysis of satellite data which confirmed the ongoing regional and seasonal nature of annual landscape fires.[6] We used of visibility and weather data to estimate PM_{10} during the study period. This has now been clarified in the Methods section

Observed PM data were only available in three years of the ten year study period - no empirical data, other than visibility and meteorology were available for seven years of the study. We chose to use one method of exposure estimation for our entire study period to be consistent throughout.

1.3. Model structure: Is this a linear model based on daily values? Were lagged values considered (e.g., the relationship between yesterday's conditions and today's PM10)? Perhaps include the final model in equation form with the coefficients used.

Yes, this is a linear model based on daily values developed using Information Theoretic methods and Gaussian linear mixed modelling. We describe the model in more detail on page 10 (14-23) and provide the model coefficients.

1.4. General questions: Why were PM_{10} data for April-June only used to estimate the values? If the model wasn't used to predict the wet season values, where did these values come from? (I found this section confusing).

Wet season periods were excluded from this analysis as 80% of Darwin's average annual rainfall (1700 mm) falls during this period, landscape fires are absent and airborne PM is consequently negligible.[7, 8] We have explained this more clearly in the text on page 7, line 6-12 following the subheading: ***Study period***

1.5. More information on model performance: The R-squared and correlation are provided, but these do not give indication of the relationship of overall values (e.g., estimated values could be consistently higher than true values and give a high r-squared). Figure 2 helps provide some evidence on this, but it would be nice to also provide summary statistics (average, median, range, etc.) on the estimated and true values. The mean deviation is provided (-2.6 microgm/m3), but it's difficult to put this into context without knowing the overall mean (i.e., is 2.6 a high or low %)?

Summary statistics have been included in this section with the addition of a new table 1: ***'Summary statistics for measured and predicted PM10 Darwin, April – June 2005'***, and referenced on page 11, line 10.

2. Given that the exposures are estimated and that multiple models were explored to estimate exposure, provide sensitivity analysis of results based on: (1) actual monitor values; and (2) alternative model structure for estimating exposure.

We did not provide this information as the differences between candidate models were so small.

3. Are estimated PM10 values available for every day? If so, the authors can investigate cumulative lag structures (e.g., average of same and previous day) rather

than single day lags. Studies in the U.S. and elsewhere have sometimes been limited to single day lags because data were not available every day.

We only examined single day lag structures.

4. Several times in the paper (e.g., abstract, Results section on page 12) the authors note that the small sample size causes large confidence intervals, which means the authors are assuming an actual association. However, the large confidence intervals could be a function of many other factors, such as uncertainty in exposure estimates, variation in population response, lack of an association, etc. It is appropriate and necessary to note the small sample size but the current wording is misleading.

These other reasons for the wide confidence intervals have been included with our discussion of the small population size on page 17 lines 4-7. We believe that our findings are consistent with other studies of ambient biomass smoke that suggest an actual association, and contribute to the limited evidence concerning the health effects of vegetation fire PM_{10}.

5. The statistically significant negative associations for COPD and indigenous people need more exploration. If the positive results are to be interpreted as the authors describe throughout the paper (positive results are indicative an association), then the statistically significant negative results would be strong evidence of a protective association. The authors need to explain why this is not the case or not make the assumption that positive results are indicative of an association. This aspect of the paper is quite problematic.

We have included additional text regarding the uncertainty in our results for the smaller sub-groups (on page 16, lines 15-17 and page 17, lines 4-5) and we have been more conservative in our interpretation. We have included text regarding the negative association with COPD on page 16, lines 18-22 and page 17, lines 1-3.

6. A key point, and strength, of this paper is the comparison of effects for indigenous and non-indigenous people, however this issue is not well explored in the current version. Please specify whether there any associations for which the effect estimates for these two groups are statistically different, and if there are none, please specify this instead.

We have expanded our descriptions of the difference in our effect estimates between indigenous and non-indigenous people. We specify that the interaction term between

indigenous status and PM_{10} was statistically different (p-value = 0.01) for respiratory infections and PM_{10} at lag 3 in the abstract (page 3 lines 1-4); in the results section (page 15, lines 11-14); and in the discussion (page 21, lines 20-21).

The interaction term also substantially improved model performance assessed using the Akaike Information Criterion (AIC) for respiratory infections and PM_{10} at lag 3. The QAIC (which adjusts the AIC for overdispersion using quasiliklihood theory) for the model with the interaction indicated that this model had substantially more support than the model without an interaction term, indicated by a difference greater than four.[9] We have not included these additional tests in the paper in order to keep the descriptions simpler.

Minor Essential Revisions

1. It would be useful to provide any quantitative estimates in the abstract.

These have been included on pages 2 and 3.

2. Please reference more of the literature on source apportionment, etc. relating to particulate matter. Currently the Laden et al. is cited (page 4), but there are many other articles including work by George Thurston, Phil Hopke, and others.

We have included additional references as suggested. Additional text regarding source apportionment has been added on page 4, lines 18-21.

3. Typo on page 5, line 9: extra period after "2000". Same after MCAPS and NMMAPS on page 9 (near bottom of page)

These have been amended.

4. Page 7. 2nd paragraph: The statement that PM_{10} is homogenous across the city needs to be removed or more evidence provided. High correlation between two monitors does not provide this evidence because: (1) such analysis is based on just 2 locations in the city; and (2) high correlation shows that the values co-vary, not that they are the same. Perhaps better wording would be that there is evidence that the monitor values for PM10 are representative of the community's exposure.

We have further justified this on page 6, lines 1-7 noting that:

1. South-easterly winds bring vegetation fire smoke over Darwin from a large region of savanna.

2. The lower atmosphere of the airshed is characteristically stable.
3. These conditions produce similar concentrations of ambient PM_{10} across the city.
4. PM_{10} measurements at two monitors located 25 km to the west and south were shown to be of similar magnitude and highly correlated with the primary monitor.

We have also included a statement similar to that recommended by the reviewer regarding evidence that the monitor values are representative of the community's exposure on page 6, line 8.

5. The statistical methods used here, developed by Peng and colleagues, addresses influenza through long-term trend. Please explain why influenza was addressed separately in this work.

The model includes both a long term trend smooth and a parameter for influenza epidemics. We included influenza epidemics in the model as local directly measured data were available. Controlling for influenza in this way is likely to be more effective than using only long term trend.

6. The text at the top of page 12 (. . . reported here as the exponent value of the model coefficient . . . multiplied . .. Subtracting one and . . .) is unnecessary and should be removed. The association should be expressed as the % increase in risk of admissions for a given increment of PM10, or as relative risk, and the formulas to calculate this do not need to be specified.

This text has been removed.

7. The paper should note that a key reason why PM_{10} effect estimates may differ by region is the different sources and resulting chemical composition of particles, such as the biomass burning noted here. There are several references relating to the chemical components of PM (mainly $PM_{2.5}$) that could be cited.

We have inserted additional text and references regarding differences in effects between sources in the discussion on page 20, lines 18-23 and page 21, lines 1-17.

***Discretionary Revisions**

1. What is the difference between this work and the Johnston et al. 2007 publication? Is it only the longer time-series and methods? Please provide more detail (briefly discussed top of page 6).

This study addresses the limited statistical power of the previous case-crossover study by assessing a longer time period and using Poisson time series modelling. Also see our response to reviewer one, major revision 3.

2. Page 6, Methods section: Does "unit record data" mean records of individual admissions?

Yes. This phrase has been replaced with 'De-identified individual records" on page 7 lines 14-16.

3. Place the figure of estimated PM_{10} concentrations (Figure 1) after explanation of the model used to generate these values.

We have moved figure 1 (now figure 2) to follow the discussion of the development of the PM predictive model and figure 2 (now figure 1, sections A and B).

4. Table 1. This table provides descriptive statistics of the resident population. However, Table 3 provides the descriptive statistics of the hospital admissions data, which is far more useful. Table 1 could be omitted.

Table 1 has been deleted. A new table replaces this table showing summary measures of the observed and predicted PM_{10} measures in response to reviewer 2 major comment 1.5 above.

5. Table 2. "Moving average" is more commonly used than "rolling average". Please specify in the table that the PM10 lagged variables are single day lags (either lag 0, lag 1, lag 2, or lag 3), rather than lag0-3 averaged. For the time variable, specify whether this is 4df per what type of season (wet/dry, season of year, fire season).

These corrections and clarifications have been done.

6. In Table 4, I believe the units for hospitalizations are of hospitalizations/day. Please specify.

This is correct. We have amended the caption to clarify

7. I appreciated the sensitivity analysis of the degrees of freedom used for the smooth function of time. However this should not be in the Methods section, but rather in the Results section.

The text describing the sensitivity analysis for the optimal smooth function of time has been moved to the results section (page 14, lines 9-19).

8. Figure 3. The y-axis labels for the first and second frames run together and are thereby confusing. Separating the frames or decreasing the font size of the y-axis labels might help.

The font size of the y-axis disease category labels and lag identifiers has been reduced to avoid them running together. This allows us to keep the 2 frames together which we believe aids comparison of the different effect estimates in the two population subgroups.

9. The numbers on Table 4 aren't in horizontal alignment with the words, making it a bit difficult to read.

The numbers in table 4 have been formatted to align with the words.

10. Sample sizes for COPD and asthma are much smaller (<50%) than the other causes. Given this limitation, the authors may consider removing these causes from the paper.

We agree with this comment, however as COPD was included in the original study design and analysis (and the results were negative) we thought it responsible to include this.

REVIEWER 3

Major Compulsory Revisions

1. State clearly in exposure measures where was the primary monitor located (in the centre of the city? Height? Relation to other pollution sources (traffic etc.)? is this a background monitor?) .

These details are now clarified in the methods section of the paper. As the topography of Darwin is flat, and 95% of measured PM arises from regional fires [8], the contribution from traffic is minimal and there are no industrial sources. Measurements taken at several locations across the city are highly correlated

regardless of proximity to the larger roads (see page 9, lines 15-23 and page 10, lines 1-3).

2. Consider leaving out Table 2 and describing the variable definition in the text instead (this depends on the Journal requirements on the number of Tables and Figures allowed).

We have retained this table and made amendments following the comments by reviewer 2 (discretionary revision 5).

3. Not appropriate to discuss the paper which is not in press yet Morgan's article from Sydney which is listed as "submitted for publication".

We have now replaced this reference with a citation of a published abstract describing this work.

Minor Essential Revisions

Sloppy and inconsistent use of punctuation! Specific suggestions below, but suggest editing article for this carefully.

We have revised our use of punctuation throughout the article and made the changes to specific words and sentences as requested.

REFERENCE LIST

1. Bowman D, Dingle JK, Johnston FH, Parry D, Foley M: **Seasonal patterns in biomass smoke pollution and the mid 20th-century transition from Aboriginal to European fire management in northern Australia**. *Glob Ecol Biogeogr* 2007, **16**:246-256.

2. Peters A, von Klot S, Berglind N, Hoermann A, Loewel H, Nyberg F, Pekkanen J, Perucci CA, Stafoggia M, Sunyer J, Tiittanen P, Forastiere F: **Comparison of different methods in analyzing short-term air pollution effects in a cohort study of susceptible individuals**. *Epidemiological Perspectives and Innovations* 2006, **3**:10.

3. Fung KY, Krewski D, Chen Y, Burnett R, Cakmak S: **Comparison of time series and case-crossover analyses of air pollution and hospital admission data**. *Int J Epidemiol* 2003, **32**:1064-1070.

4. Trewin D: **National Aboriginal and Torres Strait Islander health survey, 2004–05 (Cat. No. 4715.0)**, Canberra, Australia: Australian Bureau of Statistics, 2006.

5. Australian Bureau of Statistics: **Experimental estimates and projections, Aboriginal and Torres Strait Islander Australians, 1991-2009 (Cat. No. 3238.0).** Canberra, Australia: Australian Bureau of Statistics, 2004.

6. Russell-Smith J, Yates C, Edwards A, Allan GE, Cook GD, Cooke P, Craig R, Heath B, Smith R: **Contemporary fire regimes of northern Australia, 1997-2001: change since Aboriginal occupancy, challenges for sustainable management**. *International Journal of Wildland Fire* 2003, **12**:283-297.

7. Artaxo P, Parry D, Gillett R, Selleck P, Ayers G: **Black carbon, elemental and ionic composition of atmospheric aerosols in northern Australia**. *Global Atmosphere Watch* 1995, **107**:186-189.

8. Gras J, Parry D, Jong T, Mungksgaad N, Ayers G, Keywood M, Boast K, Powell J, Selleck P, Firestone T: **A pilot study of the air quality in Darwin NT**, Darwin, Australia: Department of Infrastructure Planning and Environment. Northern Territory Government, 2001.

9. Burnham KP, Anderson DR: **Kullback-Leibler information as a basis for strong inference in ecological studies**. *Wildlife Research* 2001, **28**:111-119.