Technical Appendix – Residential exposure to aircraft noise and hospital admissions for cardiovascular diseases: multi-airport retrospective study

Andrew W Correia, Junenette L Peters, Jonathan I Levy, Steven Melly, Francesca Dominici

Estimating Exposure to Noise:

We used noise exposure estimates (as estimated by the Integrated Noise Model, INM, described in the main text) at the census block level in combination with 2010 U.S. Census data on population counts, also at the census block level, to obtain aggregated measures of exposure to aircraft noise at the zip code level. We calculated exposure to noise metrics for 2218 zip codes around the 89 airports.

More specifically, we assumed that the study population was uniformly distributed within a census block. We then overlaid noise estimates by census block with the population ≥ 65 years of age by census block, based on U.S. Census 2010. For zip codes that included census blocks that were outside of the 45 dB contour lines provided by INM, we assumed that those blocks were exposed to 45 dB. The zip codes for which no census blocks had noise estimates above 45 dB were omitted from the analysis. We then constructed a number of candidate exposure metrics, as described in the main text, but ultimately focused on the following two: 1) population-weighted arithmetic mean noise within each zip code, weighted by the ≥ age 65 population and 2) the 90th percentile noise exposure among the census blocks within each zip code that contain non-zero population ≥ age 65. More formally, for each zip code, z, we calculated the population-weighted noise exposure by using the formula $x_z = (\sum p_j \times x_j)/h_z$, where p_j denotes the population ≥ 65 years old in census block j, x_j denotes exposure to noise at the centroid of census block j, and h_z denotes the total population ≥ 65 years old in zip code z. We summed over all census blocks (j) included in the zip code (z) separately for each z. For blocks that were split by the 45 dB contour line, noise exposure was calculated as the population exposed to over 45 dB, (estimated as the population p_j multiplied by the fraction of the block area inside the contour line f_j), multiplied by x_j, plus the estimated population outside the contour line multiplied by 45 dB. For these "split" census blocks, the block-level exposure x_j^{split} can be written as $x_j^{split} = x_j \times f_j \times p_j + 45 \times (1-f_j) \times p_j$.

Estimating Road Density as a Proxy for Road Noise and Near-Road Air Pollution:

The density of major roads within 200 m of census block centroids was estimated as a proxy for multiple risk factors correlated with roadway proximity, including traffic-related noise and air pollution. Population weighted road density within each zip code was calculated as $r_z = (\sum p_j \times r_j)/h_z$, where p_j denotes the population ≥ 65 years old in census block j, r_j denotes road density at the centroid of census block j, and h_z denotes the total population ≥ 65 years old in zip code z. Major roads were defined as limited access highways, primary roads without limited access, and secondary and connecting roads (Census Feature Class Code A1, A2 or A3) from a nationwide road dataset (StreetMap™, ESRI ArcGIS 10 Data and Maps).

Statistical Model:

We aggregated the Medicare enrollees living in areas with greater than 45 dB of airport-related noise to the zip code level, stratified by age/gender/race. Each zip code represented part of a cluster of other zip codes around one of the 89 airports. We then fit the following hierarchical Poisson model:

$$\log(E[Y_{z,s}^A]) = \log(N_{z,s}^A) + \beta_0^A + \beta_1^A I(age > 75) + \beta_2^A I(sex = M) + \beta_3^A I(race = nonwhite) + \beta_4^A(x_z - \bar{x}) + \gamma^T W_z^A$$

where $Y_{z,s}^A$ and $N_{z,s}^A$ are the number of CVD hospitalizations and number of Medicare enrollees, respectively, for zip code *z*, surrounding airport *A*, in age/gender/race stratum *s*; x_z is either the population-weighted noise exposure in zip code *z* or the 90th percentile of noise exposure among the populated census blocks that make up zip code *z*, depending on the model; and W_z^A is a vector of potentially confounding variables (e.g. $PM_{2.5}$ concentration, ozone concentration, other SES or demographic information) for zip code *z*.

For each cardiovascular outcome (combined CVD hospitalization, and separate sub-analyses for cerebrovascular disease, ischemic heart disease and heart failure hospitalizations), we constructed three models for each of the two noise exposure metrics (population-weighted noise exposure and 90th percentile of noise exposure). Models 1 controlled for individual-level variables (age, sex and race); Models 2 additionally controlled for zip code-level SES and demographic variables (% Hispanic and median household income); and Models 3 added pollution variables ($PM_{2.5}$ and ozone levels). All models were fit using R statistical software, version 2.15.2.

For our main analysis, with combined CVD hospitalizations as the outcome, we ran the hierarchical Poisson model defined above under a fully Bayesian approach,[1] allowing us to estimate the posterior distributions of both the airport-specific effects as well as overall population-level effects. The Bayesian analysis was carried out without making strong prior assumptions about the values of the regression parameters; more precisely, we used unit information priors[2 3] for the airport-specific regression parameters and flat, or uniform, priors on the fixed-effect regression parameters.

The posterior distributions obtained from this analysis can then be used to determine the posterior probability that the relative rate of CVD hospitalizations associated with airport noise exposure has a particular value. Some important calculations include the posterior mean, the posterior probability that the relative rate of CVD hospitalizations associated with airport noise is greater than zero, and the 95 percent posterior interval – a formulation similar to the 95 percent confidence interval.[4] This type of Bayesian analysis is very computationally intensive. Thus, secondary analyses that examined separately the association between noise and cerebrovascular disease hospitalizations, ischemic heart disease hospitalizations, and heart failure hospitalizations were fit using the glmer() function in the linear mixed effects package (lme4) in R, which fits the hierarchical models in a more traditional framework.[5] This method is very similar to the fully Bayesian approach used for the main analysis, and, indeed, estimates of the relative rates of total CVD hospitalization associated with noise exposure obtained using the lme4 package were nearly identical to those obtained from our fully Bayesian analyses across all models.

Population Attributable Fraction:

We calculated the population attributable fraction (PAF) using a standard expression[6]:

$$PAF = \frac{\sum_{i=1}^{n} P_i RR_i - \sum_{i=1}^{n} P_i' RR_i}{\sum_{i=1}^{n} P_i RR_i}$$

where P_i reflects the proportion of the population at exposure level i given current exposure, P_i' reflects the proportion of the population exposed to exposure level i given the counterfactual level of exposure, RR_i is the relative risk at exposure level i, and n is the number of exposure levels. For all exposures, we assumed a linear exposure-response function with a counterfactual exposure at the lowest level of exposure in the population. Exposure was treated continuously for both noise and air pollution, dividing the population into 0.01 dB, 0.01 μg/m^3, and 0.01 ppb bins for noise, $PM_{2.5}$, and ozone respectively, effectively calculating PAF as an integral as recommended for continuous exposures.

As the counterfactual level of exposure was a defined constant across the population (with a relative risk of 1), and the relative risk was applied assuming linearity throughout the range of exposures, the equation for PAF can be re-expressed as:

$$PAF = \frac{\sum_{i=1}^{n} P_i(RR_i - 1)}{\sum_{i=1}^{n} P_i(RR_i - 1) + 1}$$

All estimates were derived from a modified version of Model 3, which included additional area-level covariates that might confound the air pollution-CVD hospitalization association although they were shown to not influence the noise-CVD hospitalization association. Specifically, we additionally controlled for zip code-level high school graduation rate, black rate, urban rate, and average annual temperature. As anticipated, the noise relative rate was essentially unchanged from Model 3, with a 3.6% increase in CVD hospitalizations per 10 dB increase in 90th percentile noise (versus 3.5% in Model 3). The estimate for $PM_{2.5}$ corresponded with a 1.4% increase in CVD hospitalizations per μg/m^3 increase in annual average $PM_{2.5}$, and the estimate for ozone corresponded with a 0.56% increase in CVD hospitalizations per ppb increase in annual average ozone.

References:

1. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian data analysis*. London: Chapman & Hall, 1995.
2. Kass RE, Wasserman L. A Reference Bayesian Test for Nested Hypotheses and Its Relationship to the Schwarz Criterion. *J Am Stat Assoc* 1995;90(431):928-34.
3. Hoff PD. First Course in Bayesian Statistical Methods. *Springer Texts Stat* 2009:1-270.
4. Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994. *N Engl J Med* 2000;343(24):1742-9.
5. Bates D. Linear mixed model implementation in lme4. Department of Statistics: University of Wisconsin – Madison, 2012.
6. World Health Organization. Metrics: Population Attributable Fraction (PAF), http://www.who.int/healthinfo/global_burden_disease/metrics_paf/en/index.html.