

Department of Transportation
Federal Aviation Administration Office of Environment and Energy

SUPPORTING STATEMENT

Supporting Statement for a New Collection RE:

National Sleep Study

OMB Control Number _____

INTRODUCTION

This information collection is submitted to the Office of Management and Budget (OMB) to request a three-year approval clearance for the information collection entitled National Sleep Study, (OMB Control No. 21XX-XXXX)

Part B.

Table of Contents

<u>Part</u>		<u>Page</u>
B.1	Respondent universe and sampling methods.....	4
	B1.1 Respondent universe and sample frame.....	5
	B1.2 Sampling design.....	8
	B.1.3 Sample size determination for field study.....	8
	B.1.4 Statistical methods for field study.....	15
B.2	Procedures for the collection of information.....	16
	B.2.1 Procedures for postal survey.....	16
	B.2.2 Procedures for the field study.....	17
B.3	Methods to maximize response rates and deal with non-response.....	17
	B.3.1 Computing response and participation rates and adjusting for non-response and non-participation.....	17
	B.3.2 Maximizing response and participation rates.....	19
	B.3.2.1 Postal survey.....	19
	B.3.2.2 Field study.....	20
	B.3.3 Addressing non-response and non- participation.....	20
B.4	Test of procedures or methods to be undertaken.....	21
B.5	Individuals consulted on statistical aspects and individuals collecting and/or analyzing data.....	23
	References.....	25

Figures in this document

Figure B1 Low traffic runway groups.....	7
Figure B2 Medium traffic runway groups.....	7
Figure B3 High traffic runway groups.....	7
Figure B4 Empirical 95% confidence interval for $P(\text{Awake} L_{AS,max}) - P(\text{Awake} L_{AS,max} = 30)$ for the medium-high traffic scenario with type 1 random effects (left) and the type 3 random effects (right).....	13

Tables in this document

Table B-1 Summary of runway end classification.....	6
Table B-2 Half-width of confidence interval of $P(\text{Awake} L_{AS,max} = 50) - P(\text{Awake} L_{AS,max} = 30)$ with population based sampling and random effects variance type 1: empirical average of all 4 pilots.....	11
Table B-3 Half-width of confidence interval of $P(\text{Awake} L_{AS,max} = 50) - P(\text{Awake} L_{AS,max} = 30)$ with population based sampling and random effects variance type 2: 120% of empirical average of all 4 pilots.....	12
Table B-4 Half-width of confidence interval of $P(\text{Awake} L_{AS,max} = 50) - P(\text{Awake} L_{AS,max} = 30)$ with population based sampling and random effects variance type 3: 150% of empirical average of all 4 pilots.....	12
Table B-5 Half-width of confidence interval of $P(\text{Awake} L_{AS,max} = 50) - P(\text{Awake} L_{AS,max} = 30)$ with population-based sampling, random effects variance type 3: 150% of empirical average of all 4 pilots, and using Med-High Traffic runway class (202 Runways; 77 Airports).....	14

B.1 Respondent universe and sampling methods

Undisturbed sleep of sufficient length is of paramount importance for the maintenance of health and well-being (Watson et al., 2015). The auditory system has a watchman function and is constantly monitoring our environment for threats, including while we sleep. Noise has been shown to be a potent disruptor of sleep (Basner et al., 2014), and is considered one of the most detrimental environmental effects of air traffic (Basner, Griefahn, & van den Berg, 2010).

With the most recent US sleep study dating back to 1996 (Fidell, Pearsons, Tabachnick, & Howe, 2000), US research on the effects of aircraft noise on sleep, particularly compared to the efforts of some European countries, has lagged over the past 20 years. During the intervening time, US air traffic has changed significantly, with changes in numbers of operations on one hand, and significant reductions in noise levels of single aircraft on the other. Also, past US studies on the effects of aircraft noise on sleep predominantly used the so-called “push button” method, where study participants were required to push a button whenever they woke up during the night. This method has been shown to have low sensitivity, as most awakenings are too short for subjects to regain waking consciousness and initiate a response (Basner, Brink, & Elmenhorst, 2012). Therefore, most awakenings relevant for sleep recuperation are missed by this methodology.

Due to inter-cultural differences and different operational procedures, results from studies performed outside the US may not translate directly to US domestic airports. Therefore, it is important that field studies be conducted in the US to acquire current data on sleep disturbance relative to varying degrees of noise exposure. For this purpose, we developed and validated a methodology to unobtrusively measure noise-induced awakenings with a small device attached to the chest with only two electrodes (Basner, Griefahn, Muller, Plath, & Samel, 2007; Basner, Müller, Elmenhorst, Kluge, & Griefahn, 2008; McGuire, Müller, Plath, & Basner, 2014). The device measures body movements and heart rate, two variables strongly associated with awakenings (Basner et al., 2007). This methodology has been piloted at two US airports (Philadelphia and Atlanta) and was found to be feasible for a larger-scale national study.

The main purpose of the National Sleep Study is to collect nationally representative information on the effects of aircraft noise on sleep in order to derive exposure-response relationships between the A-weighted maximum sound pressure level $L_{AS,max}$ of single aircraft noise events (ANEs) and the probability to wake up.

The study population of the National Sleep Study is residents living close to airports, who are exposed to levels of nighttime noise from air traffic relevant for potential effects on sleep. Since airports differ in nocturnal flight operations and pattern, it will be necessary to investigate several airports across the US that are representative for all US airports with relevant nocturnal air traffic to achieve this goal. Night time aircraft noise exposure for the sampling population will be assessed using the maximum sound pressure level ($L_{AS,max}$) and the long-term energy-averaged sound pressure level during the nighttime period (L_{Night}), both expressed in decibels (dB). A-weighted average sound pressure levels (L_{Aeq}), of which L_{Night} is a special form occurring during the specified nighttime period (22:00-07:00), is considered by the US Environmental Protection Agency as the “best measure for the magnitude of environmental noise” (p. 15)(U. S. Environmental Protection Agency, 1974). According to the Federal Interagency Committee On Noise (FICON), $L_{AS,max}$ of a single flyover is useful for analyzing short-term responses (Federal Interagency Committee On Noise (FICON), 1992).

This study will conduct a field study in the homes of respondents living in the vicinity of airports that have a relevant amount of nocturnal air traffic. A sound recorder and physiological measurement equipment will be mailed to those respondents who are interested and eligible to participate in the field study (see B1.2 Sampling design). We will perform noise and physiological (heart rate, body movements) measurements in these respondents for five consecutive nights (Section **B.2.2 Procedures for the field study**).

Respondents will start and stop measurements each night themselves. We previously established the feasibility of this approach in a pilot field study around Atlanta International airport (Smith et al., 2020). Sound recordings will be used to determine the maximum sound pressure level ($L_{AS,max}$) of individual aircraft noise events in the bedroom. A previously validated method (Basner M. et al, 2007) will use heart rate and body movements to determine whether a respondent woke up in response to an aircraft noise

event. Participants will be recruited for the field study with postal surveys sent to randomly selected households exposed to a minimum level of nighttime aircraft noise (Section B 1.1).

B1.1 Respondent universe and sample frame

The respondent universe is residents living close to airports, who are exposed to levels of nighttime noise from air traffic relevant for potential effects on sleep. To determine whether residents living in certain areas around an airport were exposed to relevant levels of nighttime aircraft noise, we determined noise exposure separately for each runway. To be included in the sample frame, a runway had to meet two eligibility criteria. The first criterion was that runways had to have, on average, a minimum of one aircraft flight operation per hour during the nighttime (22:00-07:00), as determined from FAA operations data from 2018. For this purpose, both arrivals and departures that fly over a similar geographic region were counted as operations. A total of 111 airports in the contiguous United States plus Alaska, Guam, and Hawaii met this criterion. The second eligibility criterion was that airports maintain medium-high traffic during a typical sleep period, as determined by our pilot data. The 111 potentially eligible airports were classified into high, medium, and low traffic airports using the following approach. For each of the 666 runway ends at the 111 airports, we simulated our sleep study using 2018 flight operations data. For 50 out of 52 weeks (excluding the weeks of Thanksgiving and Christmas), we simulated 10 subjects living under each runway end (i.e., 500 simulations per runway end total; arrivals on a specific runway end and departures from the opposite runway end were combined as they fly over a similar geographic region). We randomly drew a subject from the Atlanta and Philadelphia pilot studies and used the observed sleep period times of that subject for our simulations, randomly drawing a sleep period time from the selected subject (with replacement) for 5 nights total. We counted the number of ANEs this simulated subject would have been exposed to during the 5 nights. Raw data and percentiles for observed number of events were stored by runway end. Runways were further divided into 3 classes based on the criteria outlined below.

- Awakening probability attributable to noise at the highest noise levels

experienced in the bedroom is typically ~10%. Several countries use 1 awakening induced by aircraft noise as a criterion for limiting the effects of aircraft noise on sleep. Thus, minimally 10 events per night (50 events per 5 nights) are needed to reach one awakening attributable to noise (i.e. 10 events each with a 10% awakening probability equals one awakening), and minimally 20 events per night (100 events per 5 nights) are needed to reach two awakenings attributable to noise. We used the median number of observed events during simulated nights for classification.

- Also, instances where investigated subjects are not exposed to a single ANE during the whole study should be rare in the actual study. Runway ends where >5% of subjects had zero observed events were classified as low traffic runways.

Based on the above criteria, runways were categorized into the following 3 classes:

- Low Traffic Runways: 0 ANEs in the 5th percentile or median number of <50 ANEs
- Medium Traffic Runways: >0 ANEs in the 5th percentile and a median number of 50-99 ANEs
- High Traffic Runways: >0 ANEs in the 5th percentile and a median number of at least 100 ANEs

The summary statistics of how many runway ends at how many different airports were assigned to each class are given in **Table B-1**. The number of simulated ANEs during the simulated sleep period in the 1st, 5th, 10th, 25th, 50th, 75th, 90th, 95th and 99th percentiles are shown in Figure B1-Figure B3 for the low, medium and high traffic runways respectively. In the field study, areas around low traffic runways will likely have too few ANEs during sleep to provide sufficient data to determine awakening probability, and so are not included in the sampling frame. There are therefore a total of 202 medium or high traffic runway ends at 77 different airports that will be included in the sampling frame and form the basis of the power calculations (see section B.1.3 Sample size determination for field study).

Table B-1 Summary of runway end classification

Runway class	Total number of runway ends across all airports	Number of airports¹
Low traffic	464 (69.7%)	34 (30.6%)
Medium traffic	119 (17.9%)	44 (39.6%)
High traffic	83 (12.5%)	33 (29.7%)
Medium or high traffic	202 (30.3%)	77 (69.4%)
Total	666	111

¹ These are the number of airports where the listed Runway class had the highest amount of flight operations, i.e. low traffic airports had only low traffic runway ends; medium traffic airports had at least one medium traffic runway end but may also have had low traffic runway ends; and high traffic airports had at least one high traffic runway end but may also have had low and/or medium traffic runway ends.

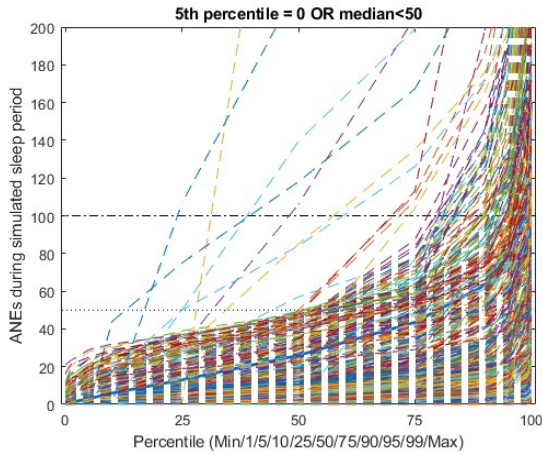


Figure B1 Low traffic runway groups

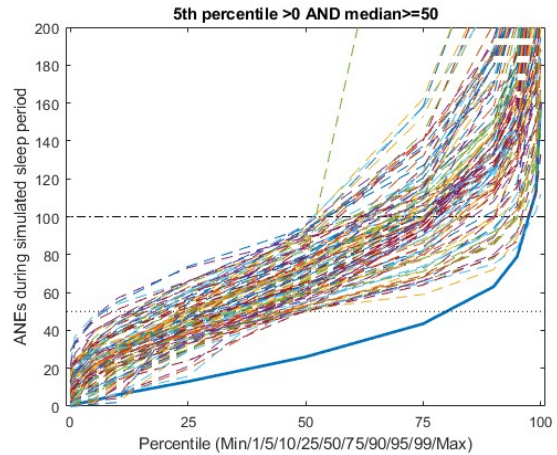


Figure B2 Medium traffic runway groups

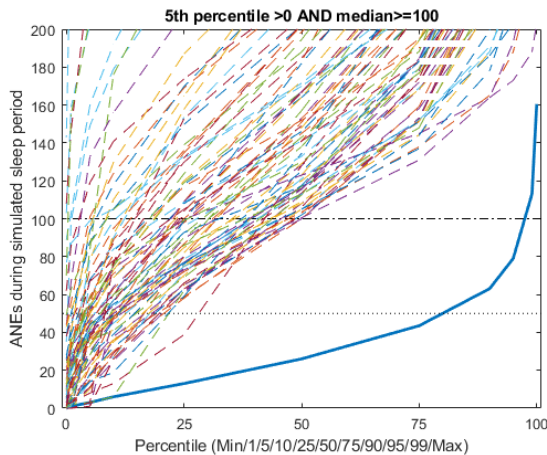


Figure B3 High traffic runway groups

Note that in each figure, each dashed, colored line represents the number of aircraft noise events (ANEs) in each percentile for a single runway end. The thick blue line represents the median number of ANEs across all runway ends at all airports, in each percentile. The horizontal dotted and dashed lines are included to indicate the median (50th percentile) cutoffs of 50 ANEs and 100 ANEs for the medium and high traffic runway classes respectively.

In addition to the number of ANEs during sleep, the sampling involves selecting addresses that are within the desired ranges of noise exposure. Prior research suggests that first reactions to aircraft noise can be expected if $L_{AS,max}$ exceeds 32-35 dB inside the bedroom (Basner, Samel, & Isermann, 2006). In the pilot study around ATL, we found these measured levels in bedrooms of households where the estimated outdoor aircraft noise level exceeded approximately 40 dB L_{night} . Also, according to the World Health Organization (WHO), “40 dB L_{night} is equivalent to the lowest observed adverse effect level (LOAEL) for night noise”. Therefore, only areas with expected aircraft noise exposure levels of ≥ 40 dB L_{night} outside will be considered for the sample frame.

For each airport selected, noise exposure contours, constructed using only runway ends classified as medium or high traffic, will be determined using the FAA's Aviation Environmental Design Tool. This will permit selection of potential survey participants by L_{night} noise exposure range and will permit computation of specific noise exposure for each sampled household. To maximize the likelihood that there will be a range of indoor noise levels in the sample frame, we will use stratified probability sampling, as described in Section B1.2 Sampling design.

In summary, the sampling frame consists of residents that live in the vicinity of runways with relevant number of expected ANEs during the nighttime (i.e., medium to high traffic runways) with an expected aircraft noise exposure level ≥ 40 dB $L_{\text{night, outside}}$ at the residential address. The residential addresses on the U.S. Postal Service Computerized Delivery Sequence File (CDSF) will be used as the household sampling frame. These addresses can be geocoded to the appropriate noise strata.

B1.2 Sampling design

The study sampling will have two phases. In the first phase, a postal survey will be sent to sampled households. In the second phase, respondents to the postal survey will be subsampled for an in-home sleep study. The purpose of this study is to develop exposure-response relationships for a potential effect of aircraft noise on awakenings.

Sampling for the postal survey will be based on a balanced stratified design with equal size strata according the following four categories of L_{night} : $40 < 45$, $45 < 50$, $50 < 55$, ≥ 55 dB. This strategy will allow for increased variability in the noise exposure for recruited participants, which increases the precision of the study. Within each of these noise strata, addresses from each of the eligible airports will be selected from the corresponding L_{night} contour region by using stratified sampling. For each noise stratum, the number of samples assigned to each airport will be determined so that the resulting sample maintains that airport's population density relative to the other airports in that stratum (determined by the number of households in that airport's L_{night} stratum relative to the total number of households in that stratum across all eligible airports, as determined by the count of addresses on the CDSF. Population-based density sampling is recommended so that the sample is

representative of the areas in the U.S. most affected by aircraft noise, that is individuals from a particular airport's affected population are sampled proportionately to the number of households affected overall in the U.S.

B.1.3 Sample size determination for field study

The primary outcome of the study is an estimated exposure-response function that describes the relationship between the maximum sound pressure level of an aircraft overflight ($L_{AS,max}$ measured inside the bedroom) and the probability of awakening (assessed with heart rate and body movements). The sample size was determined by the number needed to estimate $P(\text{Awake} | L_{AS,max} = 50 \text{ dB}) - P(\text{Awake} | L_{AS,max} = 30 \text{ dB})$ ¹ with a 95% confidence interval half-width no larger than 0.015. A half-width of the 95% confidence interval of no more than 0.015 at 50 dB $L_{AS,max}$, yielding a total width of 0.03, was considered a priori to be a just acceptable precision of the exposure-response function for regulatory purposes. Awakenings are not specific for aircraft noise, i.e., a respondent may wake up for reasons other than aircraft noise (e.g., to change body position or due to a bad dream). The quantity $P(\text{Awake} | L_{AS,max} = 50 \text{ dB}) - P(\text{Awake} | L_{AS,max} = 30 \text{ dB})$ thus describes an attributable risk, i.e., the risk of waking up to aircraft noise above and beyond the risk to wake up spontaneously (i.e., for reasons other than aircraft noise; Brink et al. 2009). A typical background noise level measured in bedrooms is 30 dB. $P(\text{Awake} | L_{AS,max} = 30 \text{ dB})$ is thus an estimate of spontaneous awakening probability (i.e., the probability to wake up without aircraft noise). The quantity $P(\text{Awake} | L_{AS,max} = 50 \text{ dB}) - P(\text{Awake} | L_{AS,max} = 30 \text{ dB})$ was chosen for the target precision because it represents the probability of awakening for an ANE in the upper range of the expected aircraft noise levels in the bedroom (approximately 75% of expected aircraft noise levels will be $< 50 \text{ dB } L_{AS,max}$) compared to a noise level that is indistinguishable from the average expected background level in the bedroom (30 dB $L_{AS,max}$, see Appendix A). The width of the 95% confidence interval for levels lower than 50 dB $L_{AS,max}$ will also be controlled; this is due to the combined factors that for increasing exposure levels, the estimated probability will become more

¹ A noise level of 30 dB $L_{AS,max}$ was based on background noise levels in the pilot studies (see Appendix A) and is also close to the noise-induced awakening threshold.

variable as it moves away from 0 (and stays less than one-half for the expected observed range of $L_{AS,max}$), and the increased uncertainty at the extreme levels of exposure of a fitted regression for the linear predictor on the logit scale (Weisberg, 2005). This required sample size was generated by a computer simulation using the following data: 1) FAA data regarding the air traffic for each airport runway to determine the number of typical aircraft events during the night (22:00-07:00) for each calendar day of the year 2018; 2) Number of ANEs expected during an individual's sleep period and corresponding $L_{AS,max}$ for those observed events based on distributions seen in our 2 pilot studies at Atlanta and Philadelphia airports, and two additional studies performed with similar methodology around Cologne and Frankfurt airports in Germany (see Error: Reference source not found for a summary of all four pilot studies); and 3) the expected exposure-response relationship, as estimated from the 4 studies (referred to as pilot studies from now on). The simulation was conducted in 4 basic steps:

Step 1: A logistic mixed effects model for probability of awakening as a function of $L_{AS,max}$ was fit to pilot data:

$$\text{Model 0: } \text{Logit}(P(\text{Awake})) = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i}) L_{AS,max}.$$

Model 0 is a generalized linear-mixed effects model in which the average exposure-response is determined by the fixed effect coefficients of a logistic regression (β_0, β_1), but allows for some variability in the intercept and slope parameters across individuals (represented by index i) by the mean-zero random intercept and slope terms (b_{0i}, b_{1i}) (Diggle, Heagerty, Liang, & Zeger, 2002). This generalized linear mixed effects model accounts for the correlation of the repeated observations within an individual. Model 0 was fit to each of the four pilot studies separately and the average of fixed effect coefficients (β_0, β_1) across the airports were used as the true fixed effect parameters for the simulation data generation, shown as Model 1 below. Details for the fitted model coefficients for the 4 pilot studies, and their average values, are provided in Appendix A.

Step 2: A logistic mixed effects model for probability of awakening as a function of $L_{AS,max}$ was used to generate hypothetical data for the National Sleep Study

$$\text{Model 1: Logit}(P(\text{Awake})) = (\beta_0 + b_{0i} + b_{0j}) + (\beta_1 + b_{1i} + b_{1j}) L_{AS,max}.$$

In Model 1 the (β_0, β_1) are the fixed intercept and slope effect coefficients. The (b_0, b_1) are the random intercept and slope effects respectively, each indexed as (i) per individual and (j) per airport. Model 1 includes airport specific random effect terms which allow for some between-airport heterogeneity in the exposure response function and the covariance between them ($\text{cov}(b_{0j}, b_{1j})$) are based on the between-airport variance of the (β_0, β_1) across the 4 pilot studies fit in Step1 (see Appendix A). Similarly, the variation of person-specific random effect terms ($\text{var}(b_{0i}, b_{1i})$) are informed by their average values across the 4 pilot models. For a given sample size, N, an airport was drawn for each individual using a multinomial distribution, where the probability of a person coming from airport j was determined by the relative population density (determined by the number of households), as described in section B.1.2. For the resulting n airports that were represented by the sample of N respondents, a set of fixed effect coefficients for Model 1 were drawn from the airport fixed effect normal distribution, where the mean and covariance matrix are set to be the average value and covariance of those parameters across the 4 pilot studies. The individual random effects were similarly drawn from a normal distribution with mean 0 and $\text{var}(b_{0i}, b_{1i})$ determined by the pilots. Once the individual exposure-response functions were determined, the number of ANEs and $L_{AS,max}$ for those events were simulated for each hypothetical individual by sampling from the pilot data. Specifically, for each simulated airport selected to be in the study, a pilot study was drawn at random and then a set of individual ANEs and $L_{AS,max}$ were sampled by sampling an individual's profile with replacement from the selected pilot study for each individual in the hypothetical data set from that airport. The awakening response was then randomly simulated according to the probability determined by Model 1 for that airport. Missing data were then simulated by removing a random 5% of individuals and then a random 30% of the observations within each given person, to match missing data rates that were seen in the previous pilot studies. The result is a fully simulated data set with \tilde{n} individuals each of whom have a varying number of events across 5 nights, with each event having an assigned $L_{AS,max}$ and awakening response.

Step 3: After data generation, the estimated exposure-response function is fit to the generated full field study data. We anticipated that a simplified version of Model 1 may need to be fit for stability purposes, in particular as the pilot studies did not suggest strong airport-specific exposure effects in the $L_{AS,max}$. Model 1 was used in the data generation process only to be conservative by building an extra source of variability for which we will still maintain the desired precision. Thus, Model 2, in which airports only have a random intercept term (b_{0j}), allowing for variability in the overall probability of awakening by airport, as well as the same two individual specific random effects (b_{0i}, b_{1i}) as before, is fit.

$$\text{Model 2: Logit}(P(\text{Awake})) = (\beta_0 + b_{0j} + b_{0i}) + (\beta_1 + b_{1i}) L_{AS,max}.$$

Model 2 is fit to the data and an estimate of $P(\text{Awake} | L_{AS,max} = 50 \text{ dB}) - P(\text{Awake} | L_{AS,max} = 30 \text{ dB})$ was obtained. In the simulation, the performance of Model 2 was compared to that if Model 1 was fit. The results were very similar between the two models and so for the purposes of the data analysis step, the simpler Model 2 was used as this is the analysis model that is anticipated to be used for the field study.

Step 4: Steps 1-3 are repeated 1000 times and the precision of the quantity $P(\text{Awake} | L_{AS,max} = 50 \text{ dB}) - P(\text{Awake} | L_{AS,max} = 30 \text{ dB})$ was determined by its empirical variation across the 1000 simulated datasets.

Table B-2 shows the results of these simulations assuming that individuals were sampled from the 77 airports with medium-high traffic as well as a second scenario with only the 33 airports with high traffic. The widths of the 95% confidence intervals between these scenarios were similar, and so the medium-high traffic scenario is preferred in order to increase the national representativeness of the airports selected. For this base scenario 300 individuals were sufficient to achieve the desired half-width for $P(\text{Awake} | L_{AS,max} = 50 \text{ dB}) - P(\text{Awake} | L_{AS,max} = 30 \text{ dB})$. However, with concerns that due to the small number of pilot studies the between-individual heterogeneity could have been underestimated, two more scenarios were run where the random effects variance terms for individuals were increased. In addition to considering the empirical average of the random effect terms across all 4 pilots (type 1 random effects variance), we also considered a 20% inflation

(type 2 random effects variance) and a 50% inflation (type 3 random effects variance). The type 3 scenario of 50% variance inflation represents a “worst case” scenario of between-individual variability. The results of these simulations are shown in Table B-3 and Table B-4. A sample size of 400 was determined to maintain the desired half-width across all scenarios. In the most variable scenario (type 3), the half-width for $P(\text{Awake} | L_{AS,max} = 50 \text{ dB}) - P(\text{Awake} | L_{AS,max} = 30 \text{ dB})$ was 0.143 for the medium and high traffic airports together and 0.149 for the high traffic airports only (Table B-4).

The numerical simulations demonstrated that a chosen sample size of 400 allows for the desired precision that is robust to some unanticipated sources of variability. If the between-person variability in the field study is closer to the empirical variance of the pilot data, then we see half width confidence intervals as tight as 0.0127 with $N=400$. We anticipate that these numbers are also conservative because the sampling design will stratify on levels of L_{night} to increase the variability in the exposure. This is expected to increase precision over that which was simulated.

Table B-2 Half-width of confidence interval of $P(\text{Awake} | L_{AS,max} = 50) - P(\text{Awake} | L_{AS,max} = 30)$ with population based sampling and random effects variance type 1: empirical average of all 4 pilots.

Runway Class	Sample Size				
	250	300	350	400	500
Med-High Traffic, Population Sampling (202 Runways; 77 Airports)	0.0153	0.0146	0.0132	0.0127	0.0111
High Traffic, Population Sampling (83 Runways; 33 Airports)	0.0147	0.0145	0.0131	0.0124	0.0119

Table B-3 Half-width of confidence interval of $P(\text{Awake} | L_{AS,max} = 50) - P(\text{Awake} | L_{AS,max} = 30)$ with population based sampling and random effects variance type 2: 120% of empirical average of all 4 pilots.

Runway Class	Sample Size				
	250	300	350	400	500
Med-High Traffic, Population Sampling (202 Runways; 77 Airports)	0.0161	0.0147	0.0142	0.0135	0.0121
High Traffic, Population Sampling (83 Runways; 33 Airports)	0.0155	0.0151	0.0140	0.0139	0.0123

Table B-4 Half-width of confidence interval of $P(\text{Awake} | L_{AS,max} = 50) - P(\text{Awake} | L_{AS,max} = 30)$ with population based sampling and random effects variance type 3: 150% of empirical average of all 4 pilots.

Runway Class	Sample Size				
	250	300	350	400	500
Med-High Traffic, Population Sampling (202 Runways; 77 Airports)	0.0168	0.0158	0.0151	0.0143	0.0129
High Traffic, Population Sampling (83 Runways; 33 Airports)	0.0176	0.0162	0.0149	0.0149	0.0131

Figure B4 shows the average and 95% confidence interval for $P(\text{Awake} | L_{AS,max} = 50 \text{ dB}) - P(\text{Awake} | L_{AS,max} = 30 \text{ dB})$ for the medium-high traffic scenario with type 1 random effects (left) and the type 3 random effects (right). As expected, the precision at lower levels of $L_{AS,max}$ are more precise than those at the target level of 50. Noise levels >60 dB in the bedroom are rare events (see Appendix A). Also in Appendix A are the exposure-response function for $P(\text{Awake} | L_{AS,max})$, with 95% confidence levels for varying levels of $L_{AS,max}$, i.e. the data used to calculate the awakening probabilities attributable to aircraft noise shown in Figure B4.

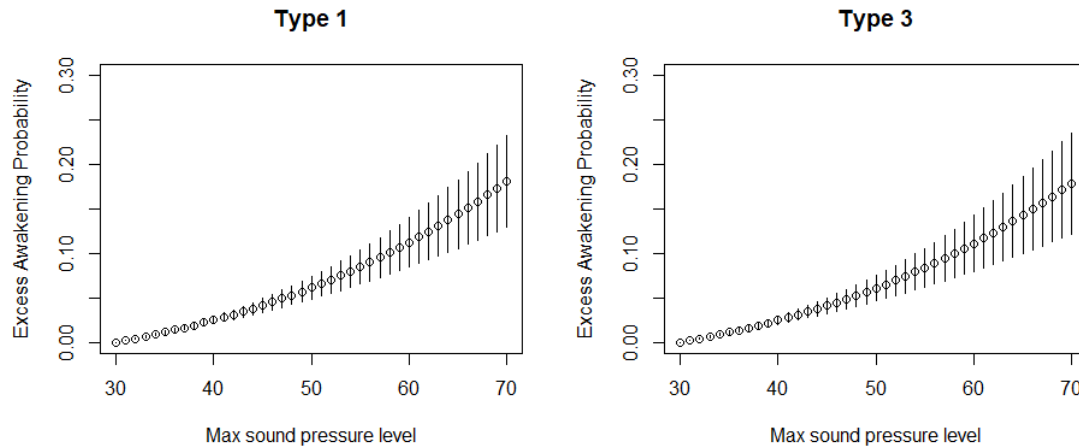


Figure B4 Empirical 95% confidence interval for $P(\text{Awake} | L_{AS,max}) - P(\text{Awake} | L_{AS,max} = 30)$ for the medium-high traffic scenario with type 1 random effects (left) and the type 3 random effects (right).

In summary, the field study will have a **sample size of N=400 individuals recruited from separate households around 77 airports throughout the U.S. Each household will have a minimum predicted outdoor aircraft noise level of 40 dB L_{night} , from nighttime aircraft operations at medium and high traffics runways.** This field study sample size forms the basis of the calculation of the number of postal surveys that will be sent to eligible households.

In response to the COVID-19 pandemic of 2020, additional simulations were done assuming varying levels of reduction in traffic across the 202 runways relative to the 2018 flight operations used to power this study. We repeated simulations that assumed the type 3 random effects variance and applied a reduction in the flight operations for all subjects, ranging from 20% to 80% (Table B-5). A sample size of 400 will maintain the desired confidence interval half-width of 0.015 for reductions at or below 30%. For a 40% reduction, 450 would be necessary to maintain the desired precision; for a 50% reduction a sample size of 500 would be necessary. After the first year of the study, an interim analysis that compares the flight operations for the 77 selected airports to those seen in 2018 will be performed in order to evaluate the relative flight operations and whether a sample size inflation is

needed to maintain the desired precision. In the unlikely event that the observed traffic drops below 50% relative to 2018 flight operations, we will consider suspending data collection. Furthermore, we will monitor the 77 airports within the sampling frame for events that could drastically change their nocturnal traffic (e.g., an airline or freight carrier discontinues using the airport as a hub), and may drop those airports from the sampling frame that no longer qualify as medium or high traffic.

Table B-5 Half-width of confidence interval of $P(\text{Awake}|L_{AS,max}=50) - P(\text{Awake}|L_{AS,max}=30)$ with population-based sampling, random effects variance type 3: 150% of empirical average of all 4 pilots, and using Med-High Traffic runway class (202 Runways; 77 Airports)

Traffic Scenario	Sample Size			
	400	450	500	550
100% traffic	0.0143	0.0134	0.0129	0.0120
20% reduced traffic	0.0144	0.0139	0.0129	0.0130
30% reduced traffic	0.0148	0.0143	0.0142	0.0128
40% reduced traffic	0.0154	0.0150	0.0140	0.0138
50% reduced traffic	0.0162	0.0153	0.0144	0.0142
60% reduced traffic	0.0166	0.0158	0.0156	0.0151
70% reduced traffic	0.0187	0.0166	0.0164	0.0152
80% reduced traffic	0.0205	0.0186	0.0181	0.0172

Based on the calculated field study sample size of 400, the overall expected sample sizes and response rates for the postal survey and participation in the field study are provided in Table B-6. The survey response rates and field study participation rates used in the table are based on a pilot study around ATL which used a similar methodology (see Section B.4 and (Smith, Witte, Rocha, & Basner, 2019)). In the event that it is necessary to recruit 450 or 500 participants into the field study, the number of postal surveys required will be increased by 12.5% or 25% respectively, in order to meet the target number of people enrolled into the field study. Also, the sampling period will be extended by 3 or 6 months for a final sample size of 450 or 500, respectively.

Table B-6 National sample sizes, response rates and expected field study enrollment

	Number
Initial sample, postal survey	24,502
12.4% PND (Postal non-deliverables) ²	3,038
20.5% response rate to postal survey, excluding PND	4,400
Field study	
58.2% of respondents are interested in field study	2,562
33.3% of interested respondents are eligible for field study	854
57.0% of eligible consent to field study participation	486
82.2% of consented are enrolled into field study	400
	%
Final postal survey response rate	18.0
Final field study participation rate, among eligible survey respondents	46.8

B.1.4 Statistical methods for field study

We will fit a generalized linear mixed effects model (Model 2) to the field study data to estimate the exposure-response function for the maximum sound pressure level ($L_{AS,max}$) of an ANE and awakening probability. Sampling weights will be used to account for the stratified sampling. Standard model diagnostics will be used to assess whether the random effects included in the model adequately captures the clustering, or whether the full Model 1 or one with fewer random effects offers a superior fit. Because the variable used for stratified sampling (L_{night}) is well-correlated with a predictor ($L_{AS,max}$) in the model, we will also consider analyses that do not use the sampling weights in the primary regression analysis. In sensitivity analyses, we may consider other weighted analyses in order to investigate the sensitivity of results to

² Postal non-deliverables are mailed surveys that were returned as non-deliverable by the USPS.

possible selection bias in the population who returned the survey versus the total that were eligible.

The effect of nighttime aircraft noise on whole-night sleep variables and morning questionnaire data will be analyzed in linear mixed models, accounting for multiple measurements on different nights of the same individuals.

B.2 Procedures for the collection of information

Postal surveys will be sent out over a 24-month period (a batch of 1/24 of all surveys each sampling month, with surveys sent continuously during 12 months of the year, making sure that each noise stratum ($40 < 45$ dB, $45 < 50$ dB, $50 < 55$ dB, ≥ 55 dB L_{night}) is equally represented in each batch. Field study participants will be investigated during 50 weeks of the year to capture seasonal effects (e.g., temperature induced changes in window-closing behavior). Field study investigations will not be performed during the weeks of Thanksgiving and Christmas as the university is closed during these periods. The scheduling of participants will always be flexible, so that they can take part whenever it is most convenient for them. Based on experience gathered during the pilot studies, we expect that the data acquisition period will last for two years, but - depending on response rates - it may last up to three years.

The data collection protocol includes two main components. The first component is a postal survey to collect data that will be used to determine if a respondent is interested in and eligible for the field study, as well as providing basic information on respondent demographics and sleep disturbance at home. The second component will be the five-night in-home field study, performed among a subsample of eligible respondents that consent to participating in the field study. The purpose of the field study is to generate exposure-response relationships between $L_{AS, \text{max}}$ of single ANEs and the probability to wake up.

B.2.1 Procedures for postal survey

The mailing protocol used for the main data collection will follow procedures outlined by (Dillman, Smyth, & Christian, 2014), and tested by us in a previous pilot study (Smith et al., 2019). All sampled addresses will be

contacted between 2 to 4 times, depending on when the questionnaire is returned. The contacts will include: 1) an initial survey packet; 2) a thank-you/reminder postcard approximately one week after the initial survey mailing; 3) a second survey package mailing two weeks after the thank-you/reminder postcard (three weeks after initial survey mailing); and 4) a third survey package mailing three weeks after the second survey package mailing.

The contents of each survey packet will include a cover letter that provides the survey purpose and sponsorship (Error: Reference source not found Appendix B for the initial mailing; Appendix C for follow-up mailings), and a paper questionnaire (Appendix D) that the respondent will be asked to return via an included postage-paid envelope or complete online. All materials mailed to the respondent will reference the 'National Sleep Study'. All survey materials will be provided in English.

A \$2 cash prepaid monetary incentive will be included with the initial mail package (See Section B.3.2.1 and (Smith et al., 2019) for rationale for incentive). The initial survey and the thank-you reminder postcard will be mailed to all sampled addresses. Only non-respondents to the prior mail packages will receive subsequent survey package mailings. Mailings returned as postal non-deliverable (PND) will be excluded from subsequent mailings.

A quasi-random selection procedure will be used to select an adult aged at least 21 years to answer the postal survey. The instructions on the first page of the survey will ask that the adult who will next celebrate a birthday should fill out the questionnaire.

B.2.2 Procedures for the field study

The postal survey packet includes a brief overview of the field study (Appendix E). If a respondent who wants to know more about the field study calls us, we will follow the script in Appendix F to provide more information and determine the caller's eligibility.

Respondents to the postal survey who indicate they are interested in the field study will be contacted by telephone if they provide their number, otherwise by email, providing they give us an email address. We will contact

only respondents who give consent for us to contact them with more information about the study (Appendix D, last page of survey). If a respondent is interested in the study but does not meet the eligibility criteria, our contact with the respondent will be to inform them as to their ineligibility. If a respondent is both interested in the study and meets the eligibility criteria, we will follow the telephone script in Appendix G to confirm the respondent's eligibility and interest, and arrange for sending them a copy of the consent form (Appendix I) to review and sign. There are minimal risks for taking part in the field study, which are described to interested respondents and explicitly stated in the consent form (pages 3-4 of Appendix I).

Upon receiving a completed consent form, we will call the prospective participant to arrange for participation in the field study, following the telephone script in Appendix H. We will then mail the participant equipment to record aircraft noise and sleep in their bedroom. This mailing will include an instruction booklet on how to set up the equipment, a single questionnaire with items to assess individual characteristics that may affect response to noise (Appendix J), and a paper questionnaire to be completed on each of the five mornings of the study (Appendix K). Once the study is complete, participants will use a FedEx phone number and shipping order number provided by us to schedule an at-home pick-up to mail the study equipment and completed questionnaires back to us. They can also return the equipment to any location that accepts FedEx shipments.

B.3 Methods to maximize response rates and deal with non-response

B.3.1 Computing response and participation rates and adjusting for non-response and non-participation

Upon completion of data collection, final response rates will be calculated for both the postal survey and field study. National estimates of the exposure-response relationship will be based on results from the field study.

For both the postal survey and field study, we will use American Association for Public Opinion Research (AAPOR) rate RR3 to compute the response rate (American Association for Public Opinion Research, 2016). For the postal survey, we will tabulate the number of postal non-deliverables for the initial

survey wave. This will provide a way to estimate eligibility among those that do not return the questionnaire:

$$RR3 = \frac{\# \text{ completed surveys}}{\# \text{ completed surveys} + \# \text{ nonrespondents} + \text{est } p(\# \text{ unknown eligibility})}$$

where #unknown eligibility is the number of sampled addresses with unknown occupancy/eligible respondent status, and *est p* is the estimated proportion of these addresses that are eligible. The proportion *e* can be estimated from the sampled addresses where occupancy/eligibility has been established. For the National Sleep Study we expect to mail to all sampled addresses so that the delivery status (and hence occupancy) of each address will be known.

All respondents in the field study sample will be a subsample of those that returned the postal survey. The postal survey includes questions for ascertaining eligibility for the field study, so we can screen out respondents who are not eligible. Therefore, we will adapt AAPOR RR1 to calculate the response rate for the field study (American Association for Public Opinion Research, 2016), which will reflect the proportion of respondents to the postal survey who are eligible for the field study and eventually participate in the field study. The number of completed surveys is here defined as the number of respondents who were both eligible for the field study and interested in taking part in the sleep study. Respondents who were not eligible or not interested are not included in the calculation. Non-respondents are defined as individuals that were eligible for the field study but who either a) we were unable to contact after receiving their completed postal survey, b) did not consent to take part in the field study, c) consented to take part but were not enrolled into the field study for any reason, or d) did not complete the full 5-day duration of the field study.

$$RR1 = \frac{\# \text{ completed surveys}}{\# \text{ completed surveys} + \# \text{ nonrespondents}}$$

$$RR1_{adapted} = \frac{i \text{ field study participants}}{i \text{ field study participants} + i \text{ eligible postal survey respondents who did not participate}}$$

B.3.2 Maximizing response and participation rates

We previously performed a pilot study to determine effective strategies to maximize response to the postal survey and maximize participation in the field study (Smith et al., 2019). These findings were used to design mail survey strategies to minimize nonresponse. The study will take proactive measures to maximize the survey response and field study participation rates.

B.3.2.1 Postal survey

The rationale for the inclusion of each question in the postal survey is given in Appendix L. These questions were chosen to maximize response rates and minimize the burden on the respondent, while providing us with necessary data to determine eligibility for the field study and perform non-response analyses. Furthermore, to maximize response to the postal survey we will take the following steps:

- **Personalized address.** The surveys will be addressed to an unnamed resident of the target area being mailed, e.g. “Philadelphia Resident”, since personalization of survey letters can increase response rates (Dillman et al., 2014).
- **Household letters.** The letters will describe the study’s sponsor, goals and objectives and will give assurances of confidentiality. Letters will be sent with each survey that is mailed to the household.
- **Multiple follow-up waves.** All sampled addresses will be sent a thank you/reminder postcard approximately one week after the initial survey mailing. If a survey is not received after the postcard, 21 days after the initial survey a second follow-up wave will be sent, with a reminder letter, a new paper copy of the survey and a new pre-paid envelope for returning the survey. If there is still no response, 42 days after the initial survey a third follow-up wave will

be sent, with a reminder letter, a further new paper copy of the survey and a further new pre-paid envelope for returning the survey. In a pilot study (Smith et al., 2019), three follow-up waves more than doubled the likelihood that a survey recipient would complete the survey, compared to when no follow-up waves were sent.

- **Use of \$2 cash incentive.** As discussed in Part A, we will include a \$2 cash incentive in the first questionnaire mailing to the household. In a pilot study (Smith et al., 2019), survey recipients sent a \$2 incentive were almost 3 times more likely to complete the survey compared to recipients offered a gift card.
- **Short survey length.** Short surveys may increase response rates (Nakash, Hutton, Jorstad-Stein, Gates, & Lamb, 2006). In a pilot study (Smith et al., 2019), survey recipients sent a 12- or 27-question survey were slightly more likely to complete the survey than recipients sent a long 58-question survey, while there were no statistically significant differences between the 12- and 27-question survey. The survey we will use consists of 25 questions.
- **Postal and online response modes.** The initial mailing round will offer only the option to respond by mail with pre-paid envelopes. In the follow-up rounds, respondents can complete the survey either by mail or online. Providing multiple response modes is an effective method to improve overall survey response (Dillman et al., 2014).

B.3.2.2 Field study

The initial step for the field study will be to check if a survey respondent is interested in participating in the field study, and then to check their eligibility for inclusion in the study. To compensate for a participant's time spent completing procedures in the field study, we will offer \$150 (\$30 per night) to individuals who participate in the 5-night in-home study. Respondents who indicate an interest in the field study will be contacted by telephone to provide them with the opportunity to ask questions and to arrange the sending of the consent form (Section B.2.2.). The respondent also has the

possibility to contact us directly. Providing the option for telephone contact ensures that potential participants are fully aware of what is required in the field study in advance, which can increase recruitment into, and reduce dropout during, the field study (Dillman et al., 2014). Telephone contact will be conducted in English only.

B.3.3 Addressing non-response and non-participation

In addition to efforts to maximize response rates, we will identify and adjust for non-response bias. There are two levels of non-response to be addressed:

1. Individuals who received the postal survey, but did not complete it
2. Individuals who completed the postal survey, but did not participate in the field study

Regarding point 1. (non-response to the postal survey), the survey includes a number of demographic questions that correspond to questions from the US Census (United States Census Bureau, 2010), the American Community Survey (United States Census Bureau, 2019) and the Behavioral Risk Factor Surveillance System (Center for Disease Control and Prevention, 2019). These questions are described in detail in Appendix L, and allow us to compare the demographics of respondents with the expected demographics in the census tracts from which we sampled. Questions on race and ethnicity comply with standards described in OMB Directive 15.

We will determine whether the respondents are representative of the sampled population (Thabane et al., 2013). If there is evidence that the survey response data are from a population not representative of the broader population, we will adjust for response bias using the most appropriate statistical methods for the data, conceivably with inverse probability weighting (Mansournia & Altman, 2016).

Regarding point 2. (non-participation in the field study), the postal survey includes questions specifically included to allow analysis of non-participation in the field study among the survey respondents (see Appendix L). We will perform statistical testing to determine if there are differences in the demographics of participants and non-participants, among survey respondents who are eligible for the field study. Based on results from the pilot study (Smith et al., 2019), we do not anticipate there will be many

differences in demographics of the field study participants and non-participants. If there is evidence for a selection bias among the participants, that is not accounted for by eligibility criteria, we will consider appropriate weighting of the responses.

B.4 Test of procedures or methods to be undertaken

The questionnaires and data collection procedures used in this study are based on a pilot study (Smith et al., 2019) conducted through ASCENT – the FAA Center of Excellence for Alternative Jet Fuels and Environment the National Academy of Sciences (ASCENT project 017; Pilot study on aircraft noise and sleep disturbance, Final report, available at ascent.aero). The purpose of this pilot study was to validate the methodology of a study to collect and analyze data for estimating the exposure-response relationship relating noise to sleep disturbance, in particular the quality and quantity of data that could be obtained when recruiting participants by postal questionnaire, shipping them the physiological and noise measurement equipment, and the setup of the equipment and recording of data by the participants themselves, unattended. This study was conducted in communities around Atlanta International Airport in the following two stages:

1. A postal survey using an address-based sample frame (ABS). An ABS frame was used because of the need to map households within specific noise strata surrounding the airport. Different postal survey strategies were used to collect the data, to determine how we could maximize response rates (see (Smith et al., 2019) for description and analysis of the different strategies).
2. A 5-night in-home study of sleep in a subset of respondents to the postal survey. This study consisted of mailing the noise and sleep recording equipment to study participants, who setup the equipment themselves, and returned the equipment and recorded data stored thereon via mail upon study completion.

The questionnaires in stage 1. were developed after an extensive review of the literature on airport noise and its relationship to sleep disturbance. The study found that personalizing the address, enclosing a \$2 cash incentive with the initial questionnaire mailing and repeated follow-up mailings were effective at increasing response rate. The likelihood that a respondent would

participate in the field study in stage 2. was unaffected by survey incentive, survey length, number of follow-up waves, field study incentive, age or sex.

Based on these results, a national study was designed that relies on postal questionnaires to determine eligibility and recruit for an in-home study to collect the data necessary to estimate a physiologic exposure-response relationship. The questionnaires that we will use in the national study, and the rationale for including them, are described in Appendix L. We expect the response rate for the national postal survey will be enhanced by several factors discussed above in Section B.3.

The primary outcome of the study is an exposure-response function between the maximum sound pressure level ($L_{AS,max}$) of an ANE and awakening probability. Aircraft noise events will be identified by human scorers listening to the bedroom sound recordings and where possible with the help of flight track surveillance data. The beginning and end of each aircraft noise event will be marked and several other acoustical descriptors will be calculated. The primary outcome of awakening (a binary yes or no outcome) will be identified using an algorithm based on the collected physiological data and screened for during a 50-second window starting 5 seconds before the start of the aircraft noise event. As outlined in detail in B.1.4 above, we will fit a generalized linear mixed effects model to the field study data to estimate the exposure-response function, and perform sensitivity analyses to investigate the robustness of the findings.

In secondary analyses, we will investigate the effect of nighttime aircraft noise on whole-night sleep variables (as measured from the collected physiological and acoustic variables) and morning questionnaire data with linear mixed models accounting for multiple measurements on different nights of the same individuals. We will also consider models relating predicted average aircraft noise levels with postal survey responses on self-reported sleep, health, and annoyance adjusting for sociodemographic variables and other potential confounders.

The study protocol has been reviewed and approved by and will be performed under the oversight of the Institutional Review Boards of the University of Pennsylvania and Westat. Details on how survey data, acoustic

data and physiologic data will be managed and stored assuring confidentiality can be found in section 10 of Part A.

At the conclusion of this study, the University of Pennsylvania will deliver a final report outlining the results of the analyses outlined above to FAA. Results will be reported in aggregate form without identifying individual study participants. The University of Pennsylvania shall maintain all study information retained in accordance with National Archives and Records Administration (NARA) records retention policies and schedules and FAA policies.

B.5 Individuals consulted on statistical aspects and individuals collecting and/or analyzing data

FAA project lead and other technical lead

Sean Doyle

Environmental Protection Specialist

Federal Aviation Administration

Mathias Basner

Principal Investigator

Professor of Psychiatry

University of Pennsylvania

Pamela Shaw

Associate Professor of Biostatistics

University of Pennsylvania

Michael Smith

Postdoctoral researcher

University of Pennsylvania

Grace Choi

Statistical analyst

University of Pennsylvania

Eric Jodts

Associate Director

Westat

Hanna Popick

Senior Study Director

Westat

Jennifer Kali

Senior Statistician

Westat

References

- American Association for Public Opinion Research. (2016). *Standard Definitions: Final dispositions of case codes and outcome rates for surveys* (9th ed.): AAPOR.
- Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S., & Stansfeld, S. (2014). Auditory and non-auditory effects of noise on health. *Lancet*, 383(9925), 1325-1332. doi:10.1016/S0140-6736(13)61613-X
- Basner, M., Brink, M., & Elmenhorst, E. M. (2012). Critical appraisal of methods for the assessment of noise effects on sleep. *Noise Health*, 14(61), 321-329. doi:10.4103/1463-1741.104902
- Basner, M., Griefahn, B., Müller, U., Plath, G., & Samel, A. (2007). An ECG-based algorithm for the automatic identification of autonomic activations associated with cortical arousal. *Sleep*, 30(10), 1349-1361.
- Basner, M., Griefahn, B., & van den Berg, M. (2010). Aircraft noise effects on sleep: mechanisms, mitigation and research needs. *Noise Health*, 12(47), 95-109. doi:10.4103/1463-1741.63210
- Basner, M., Müller, U., Elmenhorst, E. M., Kluge, G., & Griefahn, B. (2008). Aircraft noise effects on sleep: a systematic comparison of EEG awakenings and automatically detected cardiac activations. *Physiol Meas*, 29(9), 1089-1103. doi:S0967-3334(08)79908-8 [pii] 10.1088/0967-3334/29/9/007
- Basner, M., Samel, A., & Isermann, U. (2006). Aircraft noise effects on sleep: application of the results of a large polysomnographic field study. *J Acoust Soc Am*, 119(5 Pt 1), 2772-2784.
- Brink, M., Basner, M., Schierz, C., Spreng, M., Scheuch, K., Bauer, G., Stahel, W.A.: Determining physiological reaction probabilities to noise events during sleep. *Somnologie* 13(4), 236-243, 2009.
- Center for Disease Control and Prevention. (2019). Behavioral Risk Factor Surveillance System. Retrieved from <https://www.cdc.gov/brfss/about/index.htm>
- Diggle, P. J., Heagerty, P., Liang, K.-Y., & Zeger, S. L. (2002). *The analysis of longitudinal data* (2nd ed.). Oxford: Oxford University Press.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys : the tailored design method* (4th edition. ed.). Hoboken: Wiley.
- Federal Interagency Committee On Noise (FICON). (1992). *Federal agency review of selected airport noise analysis issues*. Retrieved from
- Fidell, S., Pearsons, K., Tabachnick, B. G., & Howe, R. (2000). Effects on sleep disturbance of changes in aircraft noise near three airports. *Journal of the Acoustical Society of America*, 107(5), 2535-2547. doi:Doi 10.1121/1.428641
- Mansournia, M. A., & Altman, D. G. (2016). Inverse probability weighting. *Bmj-British Medical Journal*, 352. doi:ARTN i189

10.1136/bmj.i189

- McGuire, S., Müller, U., Plath, G., & Basner, M. (2014). *Refinement and validation of an ECG based algorithm for detecting awakenings*. Paper presented at the 11th International Congress on Noise as a Public Health Problem (ICBEN), Nara, Japan.
- Nakash, R. A., Hutton, J. L., Jorstad-Stein, E. C., Gates, S., & Lamb, S. E. (2006). Maximising response to postal questionnaires--a systematic review of randomised trials in health research. *Bmc Medical Research Methodology*, *6*, 5. doi:10.1186/1471-2288-6-5
- Smith, M. G., Witte, M., Rocha, S., & Basner, M. (2019). Effectiveness of incentives and follow-up on increasing survey response rates and participation in field studies. *Bmc Medical Research Methodology*, *19*(1). doi:ARTN 23010.1186/s12874-019-0868-8
- Smith, M.G., Rocha, S., Witte, M., Basner, M. (2020): On the feasibility of measuring physiologic and self-reported sleep disturbance by aircraft noise on a national scale: A pilot study around Atlanta airport. *Science of the Total Environment*, *718*, 137368, 1-12, 2020.
- Thabane, L., Mbuagbaw, L., Zhang, S. Y., Samaan, Z., Marcucci, M., Ye, C. L., . . . Goldsmith, C. H. (2013). A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *Bmc Medical Research Methodology*, *13*. doi:Artn 92 10.1186/1471-2288-13-92
- U. S. Environmental Protection Agency. (1974). *Information on levels of environmental noise requisite to protect public health and welfare with an adequate margin of safety*. Retrieved from
- United States Census Bureau. (2010). 2010 Decennial census of population and housing. Retrieved from <https://www.census.gov/programs-surveys/decennial-census/decade.2010.html>
- United States Census Bureau. (2019). American Community Survey (ACS). Retrieved from <https://www.census.gov/programs-surveys/acs>
- Watson, N. F., Badr, M. S., Belenky, G., Bliwise, D. L., Buxton, O. M., Buysse, D., . . . Tasali, E. (2015). Recommended Amount of Sleep for a Healthy Adult: A Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society. *Sleep*, *38*(6), 843-844. doi:10.5665/sleep.4716
- Weisberg, S. (2005). *Applied linear regression* (3rd ed.). Hoboken, N.J.: Wiley-Interscience.