Bias from Differential Exposure Measurement Error in a Study of Flight Attendants

Candice Y. Johnson; Barbara Grajewski

BACKGROUND: Self-reported occupational exposures are often used in epidemiological studies when actual exposure measurements are unavailable, which could cause measurement error and bias study results. This study provides a numeric example of this potential bias.

METHODS: A study of block hours and preterm birth was used as an illustrative example. This study included 577 flight attendants, ages 18-45 yr, who gave birth to a term (37 or greater gestational weeks) or preterm (20-36 gestational weeks) infant between 1992 and 1996. Flight attendants self-reported the number of block hours flown during the first trimester of pregnancy; the number of block hours flown during the first trimester of pregnancy was also calculated from airline records. No adjustment for confounding was performed for this illustrative example.

RESULTS: Although flight attendants having term and preterm births self-reported similar hours worked during the first trimester (median 213 vs. 215 block hours), airline records showed that flight attendants having term births worked more hours than those having preterm births (median 146 vs. 104 block hours). Using self-reported block hours, there was no association between block hours and preterm birth; when using airline records, an inverse association was observed.

DISCUSSION: In this example, differential measurement error from use of self-reported block hours obscured an inverse association apparent when using airline records, demonstrating the importance of accurate exposure assessment for identifying occupational risk factors for health outcomes.

KEYWORDS: block hours, exposure assessment, preterm birth.

Johnson CY, Grajewski B. *Bias from differential exposure measurement error in a study of flight attendants.* Aerosp Med Hum Perform. 2015; 86(11):990–993.

In many epidemiological studies, researchers must rely on self-reported exposures when actual exposure measurements are unavailable. If exposures are not accurately reported by study participants, nondifferential measurement error (when error occurs independently of other variables) or differential measurement error (when error is related to another variable, such as the outcome) might occur, leading to biased study results.[5] Bias analysis can be used to adjust study results for measurement error (error on a continuous variable) or misclassification (error on a categorical variable), although validation data must be available to inform the choice of the parameters needed to perform the adjustment, which is seldom the case.[4]

When studying health outcomes among flight crew, number of block hours worked (taxi time plus flight time) can be used as one measurement of the flight crew's work hours. Previously, it has been shown that flight attendants over-report the number of block hours they work compared to the hours recorded by their airline.[1] The ways in which this error in reporting of block hours might affect the results of an epidemiological study among flight attendants have not yet been explored in detail.

This analysis provides an illustrative example of how measurement error can affect interpretation of results from an epidemiological study. The example comes from a study of block hours and preterm birth. However, this analysis is meant to illustrate the effects of measurement error on study results and not intended to determine the relationship, if any, between flight attendant exposures and preterm birth.

From the National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Cincinnati, OH, and the Epidemic Intelligence Service, Centers for Disease Control and Prevention, Atlanta, GA.

This manuscript was received for review in March 2015. It was accepted for publication in August 2015.

Address correspondence to: Candice Y. Johnson, Ph.D., National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, 1090 Tusculum Ave., MS R-15, Cincinnati, OH 45226; cyjohnson@cdc.gov.

DOI: 10.3357/AMHP.4321.2015

METHODS

Data were used from a study of reproductive health among female flight attendants.[2] The study protocol was approved in advance by the Institutional Review Board of the National Institute for Occupational Safety and Health. Each participant provided informed consent before participating. Married flight attendants ages 18-45 were selected from three airlines with hubs in Detroit, Miami, and Seattle. Flight attendants working for at least 1 mo during the study period, August 1992 to July 1996, were eligible if they had not had a hysterectomy before the start of the study period. Individual flight records were obtained from the three airlines for all flights worked by the flight attendants during the study period. Between November 1999 and April 2001, flight attendants were contacted to complete a computer-assisted telephone interview, which included questions on demographics, occupational exposures, and health outcomes.

Preterm birth was one of the outcomes of interest. In the interview, the flight attendant reported the gestational length of each of her pregnancies. Preterm birth was defined as a live born infant delivered at 20 to 36 gestational weeks, as counted from the last menstrual period. Term birth was defined as a live born infant delivered at 37 or more gestational weeks. Eligible pregnancies were those in which the first trimester occurred at some point during the study period. If a flight attendant had more than one live birth during the study period, one of these births was randomly selected for inclusion. Flight attendants who had a pregnancy lasting less than 20 wk, who had a stillbirth, or who were not pregnant during the study period were excluded from the analysis.

Block hours worked during the first trimester of pregnancy was the exposure of interest. First trimester exposure was chosen to avoid bias that results from using second or third trimester exposures in studies of preterm birth: if a woman's pregnancy ends earlier (preterm) she will have less opportunity for exposure than if her pregnancy had continued to term (e.g., if the pregnancy ended before the third trimester, the block hours worked in the third trimester would be zero). The purpose of this study was to compare results of the study when using estimates of block hours worked in the first trimester of pregnancy obtained from airline records (assumed to be accurate) to estimates obtained by self-report (assumed to be subject to measurement error). From airline records, the number of block hours flown during the first trimester of pregnancy (from date of last menstrual period through gestational week 12) was calculated directly from the flight records for each flight attendant, which included the number of block hours for each of her flights. For the self-reported estimate of block hours worked in the first trimester, because the original study was not designed for this purpose, the interview did not include a question on first trimester block hours. Instead, this was estimated from the question: "During this pregnancy, how many block hours did you usually fly per month as a requirement of your job?" The answer to this question was multiplied by 3 to obtain an estimate for the block hours worked during the first trimester of pregnancy.

Based on prior results from this cohort which suggest that flight attendants who do not fly at all during pregnancy differ in pregnancy outcome from flight attendants who fly during pregnancy, 21 flight attendants who had no first trimester block hours recorded by the airlines were excluded.[2] T-tests were used to compare mean block hours of flight attendants having preterm to those having term births, paired *t*-tests were used to compare measured to self-reported block hours, and log-binomial regression was used to estimate risk ratios (RR) and 95% confidence intervals (CI) for associations between block hours (as a continuous variable and dichotomized at the median) and preterm birth using SAS version 9.3 (SAS Institute, Cary, NC). Use of tertiles or quartiles for categorization of block hours was not possible because the small number of preterm births in the study population made these estimates unstable.

RESULTS

Of the 557 live births included in this study, 27 (5%) were preterm. Flight attendants self-reported a greater number of hours worked during the first trimester of pregnancy compared to their airline records [mean 213 self-reported block hours (SD = 72), 144 recorded block hours (SD = 65), $t(556) = -20.5$, $P < 0.0001$]. Flight attendants having preterm and term births self-reported a similar number of hours worked during the first trimester [mean 215 h for preterm births (SD = 66), 213 h for term births (SD = 72), $t(555) = -0.15$, $P = 0.88$, **Fig. 1**]. However, according to airline records, flight attendants who eventually had a preterm birth worked fewer hours in the first trimester than flight attendants who had a term birth [mean 104 h for preterm births (SD = 70), 146 h for term births (SD = 64), $t(555) = 3.36$, $P = 0.0008$].

Differential measurement error was observed when using both the continuous and dichotomized versions of the variable. For the continuous variable, the distribution of the magnitude of error (difference between measured and self-reported block hours) had a mean of −66.5 h for term births and −111.4 h for preterm births. For the dichotomized variable, sensitivity and specificity were 0.56 and 0.62 for term births and 0.47 and 0.63 for preterm births. For both variables, measurement error was greater for flight attendants who had preterm births than for those who had term births.

When using self-reported block hours, there was no association between block hours and preterm birth when using block hours dichotomized at the median (RR 1.09, 95% CI: 0.52–2.29). When using block hours estimated from airline records, an inverse association was found between dichotomized block hours and preterm (RR 0.42, 95% CI 0.19–0.94). Based on this information, it appears that misclassification biased the results toward the null. The continuous block hour variable was also used in the analysis; however, the measurement error was less clearly observed because the magnitude of the association was small when estimating relative risk per block hour of flight; the *P*-values more clearly demonstrated the difference (self-report: RR 1.00, 95% CI: 1.00–1.01, $P = 0.88$; continuous: RR 0.99,

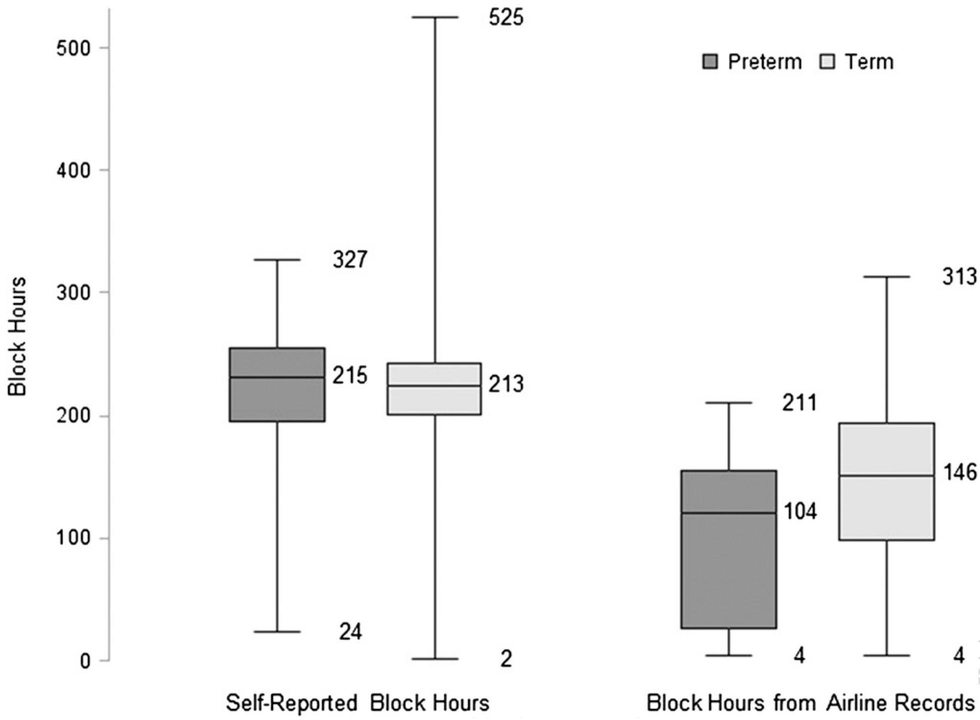

Fig. 1. Box plots showing number of block hours worked during the first trimester from airline records and self-reported in the questionnaire among flight attendants having preterm (dark gray) and term (light gray) births. Bottom and top edges of the boxes indicate the 25th and 75th percentiles of the distribution, the horizontal line in the box represents the mean, and the whisker edges represent minimum and maximum values.

95% CI: 0.99–1.00, $P = 0.001$). These associations are in the same directions as for the dichotomized variable and also suggest bias toward the null.

DISCUSSION

In this study population, there was evidence of differential measurement error for self-reported block hours according to preterm birth status, with flight attendants having a preterm birth over-reporting their block hours to a greater extent than flight attendants having a term birth. This error resulted in bias toward the null in the analysis using both continuous and dichotomous exposure variables, with a null association observed when using the self-reported estimate and an inverse association observed when using block hours from airline records. A null association between block hours and preterm birth might not have been questioned if more accurate exposure data were unavailable and the potential significance of an inverse association would have been missed. One might hypothesize that the inverse association occurs because women who have had a prior preterm birth (and who are at higher risk of future preterm birth) alter their work schedules in subsequent pregnancies to avoid another adverse pregnancy outcome.

The over-estimation of self-reported block hours observed in this study is similar to a previous study which found that flight attendants tend to over-report both their block hours and flight segments when compared to airline records.[1] Because block hours and flight segments can be used to estimate other exposure metrics, such as cosmic ionizing radiation dose or time zones crossed, use of self-reported data might result in inaccurate estimation of these metrics as well.[1] If exposure misclassification or measurement error is suspected, bias analysis can be used to adjust study results for this error.[4] However, conducting a bias analysis requires making assumptions such as whether the error was nondifferential or differential. If incorrect assumptions are made, bias analyses might produce results that do not reflect what the association would have been in the absence of exposure measurement error or misclassification.[3] Prior to this analysis, there was no evidence in the literature showing that the error on self-reported block hours would be differential by preterm status. Had a bias analysis been conducted, nondifferential error might have incorrectly been assumed and the bias analysis might not have correctly determined the magnitude and direction of the association in the absence of exposure measurement error or misclassification.

Although an epidemiological study of block hours and preterm birth was conducted to illustrate the potential for bias from exposure measurement error, the purpose of this analysis was not to estimate the effect of block hours on preterm birth. For simplicity, there was no adjustment for potential confounders, nor were potential interactions or mediators of the association explored. The results should, therefore, not be interpreted as estimates of unbiased associations between block hours and preterm birth.

Use of this population for this illustrative example had several limitations. First, in the interview, women were not directly asked about block hours worked during the first trimester, and the question on average number of hours worked per month during the pregnancy does not allow a precise determination of how reducing work hours or stopping work during the pregnancy would have figured into women's estimation of hours worked per month. The most likely result would be an underestimation of the number of self-reported block hours worked in the first trimester; however, self-reported block hours was markedly higher than block hours recorded in airline records. As a result, it might be expected that bias would be even greater had a more precise measure of self-reported block hours been used. Second, the low prevalence of preterm birth in this population compared to the general population (5% versus 11%)[6] resulted in low power and unstable estimates. It is difficult to separate the effects of bias from random error unless sample size is sufficiently large to realistically exclude random error as a possibility. Despite the small sample size of this study, differences between the term and preterm groups were large enough to produce small *P*-values, providing encouraging evidence that the differences seen are attributable to bias and not random error, and that similar results might be seen in a study with larger sample size. Small sample size also prevented exploration of further ways to categorize the continuous variable, such as using tertiles or quartiles.

This study illustrates how measurement error can distort the results of an epidemiological study and demonstrates the importance of accurate exposure assessment. The inverse association between block hours and preterm birth in this study population would not have been observed had the self-reported exposures been relied upon. Measurement error can alter the magnitude and direction of effect estimates from epidemiological studies, implying that policy decisions can be made based on studies whose results would have been different in the absence of measurement error. Studies like this one, in which airline records were used, are more resource-intensive, but improve the ability to accurately identify occupational risk factors for health outcomes.

ACKNOWLEDGMENTS

The findings and conclusions of this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention. This work was part of the Intramural Research Program of the National Institute for Occupational Safety and Health (NIOSH). The study was supported, in part, by interagency agreements with the Federal Aviation Administration (FAA) and the Department of Defense Women's Health Research Program. We thank the women who participated in this study. We appreciate the assistance of the study airlines, Airlines for America, the Association of Flight Attendants and its Seattle Local 19, the Association of Professional Flight Attendants and its Miami representatives, and the International Brotherhood of Teamsters, Airline Division, and its Detroit representatives.

Authors and affiliation: Candice Y. Johnson, Ph.D., M.Sc., and Barbara Grajewski, Ph.D., M.S., National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Cincinnati, OH.

REFERENCES

1. Grajewski B, Atkins DJ, Whelan EA. Self-reported flight hours vs. company records for epidemiologic studies of flight attendants. Aviat Space Environ Med. 2004; 75(9):806–810.
2. Grajewski B, Whelan EA, Lawson CC, Hein MJ, Waters MA, et al. Miscarriage among flight attendants. Epidemiology. 2015; 26(2):192–203.
3. Johnson CY, Flanders WD, Strickland MJ, Honein MA, Howards PP. Potential sensitivity of bias analysis results to incorrect assumptions of nondifferential or differential binary exposure misclassification. Epidemiology. 2014; 25(6):902–909.
4. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. New York: Springer; 2009.
5. Rothman KJ, Greenland S, Lash TL. Validity in epidemiologic studies. In: Rothman KJ, Greenland S, Lash TL, editors. Modern epidemiology, 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008.
6. Ventura SJ, Martin JA, Curtin SC, Mathews TJ. Report of final natality statistics, 1996. Mon Vital Stat Rep. 1998; 46(11, Suppl.):1–99.