Annals of Epidemiology xxx (2016) 1–8

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Original article

Quantitative bias analysis in an asthma study of rescue-recovery workers and volunteers from the 9/11 World Trade Center attacks

Anne M. Jurek PhD *, George Maldonado PhD

Division of Environmental Health Sciences, University of Minnesota School of Public Health, Minneapolis

ARTICLE INFO

Article history:
Received 13 May 2016
Accepted 15 September 2016
Available online xxx

Keywords:
Asthma
Bias analysis
Outcome misclassification
Selection bias
Sensitivity analysis
World Trade Center
9/11

ABSTRACT

Purpose: When learning bias analysis, epidemiologists are taught to quantitatively adjust for multiple biases by correcting study results in the reverse order of the error sequence. To understand the error sequence for a particular study, one must carefully examine the health study's epidemiologic data-generating process. In this article, we describe the unique data-generating process of a man-made disaster epidemiologic study.

Methods: We described the data-generating process and conducted a bias analysis for a study associating September 11, 2001 dust cloud exposure and self-reported newly physician-diagnosed asthma among rescue-recovery workers and volunteers. We adjusted an odds ratio (OR) estimate for the combined effect of missing data, outcome misclassification, and nonparticipation.

Results: Under our assumptions about systematic error, the ORs adjusted for all three biases ranged from 1.33 to 3.84. Most of the adjusted estimates were greater than the observed OR of 1.77 and were outside the 95% confidence limits (1.55, 2.01).

Conclusions: Man-made disasters present some situations that are not observed in other areas of epidemiology. Future epidemiologic studies of disasters could benefit from a proactive approach that focuses on the technical aspect of data collection and gathers information on bias parameters to provide more meaningful interpretations of results.

Introduction

Bias (systematic error) is a nonrandom distortion in epidemiologic study estimates. Most epidemiologic study results contain multiple sources of bias. Epidemiologists typically evaluate biases one-at-a-time (qualitatively or quantitatively); yet the combined effect of multiple biases should be evaluated. Bias analysis is a method that quantitatively evaluates the impact of study errors on study results. Bias analysis includes both nonprobabilistic and probabilistic sensitivity analyses [1–4]. To quantitatively adjust the impact of multiple biases, one should correct for each bias in the reverse order of the error sequence [1,2]. To understand the error sequence for a particular study, one must carefully examine the health study's data-generating process.

The data-generating process depends on the research question and the target population. It also requires understanding data collection, timing of human subject consent, and errors in the data. While this process is well understood for many types of epidemiologic studies (e.g., cancer case-control studies), the data-generating process for epidemiologic research on disasters is not.

In this teaching article, we describe the data-generating process and conduct a bias analysis for a man-made disaster study associating September 11, 2001 (9/11) dust cloud exposure and newly physician-diagnosed asthma among rescue-recovery workers and volunteers.

Materials and methods

Man-made disaster epidemiology

Man-made disasters present some situations that are not observed in other areas of epidemiology. The unplanned nature of man-made disasters causes the epidemiologic process to be reactive. Priority and focus are immediately given to national security, medical treatment of injured persons, and site cleanup. Scientific efforts are secondary, resulting in data collection efforts occurring after-the-fact. Additionally, data elements can be nonspecific and wide ranging. For instance, contamination and debris exposures are difficult to identify and quantify because environmental monitoring

* Corresponding author. Health Services, The Dow Chemical Company, Employee Development Center, Midland, MI 48674. Tel.: 989-638-4682; fax: 989-696-7431.
E-mail address: jure0007@umn.edu (A.M. Jurek).

http://dx.doi.org/10.1016/j.annepidem.2016.09.002
1047-2797/

data are often unavailable. Moreover, disasters such as the 9/11 attacks cause an array of outcomes including deaths, injuries, and psychological and physical illnesses. These outcomes present in workers, building occupants, passersby, residents, and volunteers. The total number of persons affected is poorly documented and difficult to enumerate (especially for transient workers [5]). Furthermore, a diverse cohort emerges when numerous people are affected that have few commonalities besides the disaster [6].

WTC 9/11 attacks

The two airplane-hijacking attacks on the WTC twin towers in New York City on September 11, 2001 was a man-made disaster. The destruction of structures and subsequent fire exposed workers and volunteers to dust and smoke containing various environmental agents and airborne particulate matter [7]. It is estimated that over 90,000 people including rescue workers, volunteers, and construction contractors assisted with rescue and recovery [8]. Cleanup was complete in June 2002, 9 months after the attack. Immediately following the attack, however, medical conditions, such as injuries and respiratory outcomes, were already present among workers and volunteers [7]. Therefore, health officials found it imperative to begin monitoring the well-being of those exposed to the WTC disaster [9].

WTC Health Registry

In response to the WTC 9/11 disaster, the Agency for Toxic Substances and Disease Registry and the New York City Department of Health and Mental Hygiene collaboratively established the World Trade Center Health Registry (WTCHR) to periodically monitor the mental and physical health of workers and survivors having direct exposure to the 9/11 attacks and aftermath. The WTCHR has been discussed in detail previously [7,9]. Briefly, created in July 2002, the WTCHR focuses on short- and long-term health effects. Data collected include 9/11 exposures, demographics, and mental and physical health outcomes. The WTCHR is the largest U.S. post-disaster registry with over 71,000 people consenting and completing the baseline interview within three years after 9/11 [7,10]. The WTCHR collected information about exposure to the dust cloud and information on health outcomes that occurred after 9/11. Eligible participants were targeted into one of four broad exposure groups: (1) workers, (2) residents, (3) students and staff, and (4) occupants and passersby. Long-term follow-up of participants continues. In 2015, the third follow-up survey was sent to enrollees [11].

Application

Using WTCHR exposure and outcome data from participants completing the baseline interview between 2003 and 2004, Wheeler et al. [12] assessed asthma among rescue-recovery workers and volunteers from the 9/11 disaster. They found a significantly increased risk of self-reported newly diagnosed (by physician or other health professional after 9/11) asthma among those exposed to the 9/11 dust cloud (more specifically, the cloud of dust or debris; unadjusted odds ratio (OR) = 1.77, 95% confidence limits 1.55, 2.01).

The authors discussed the impact of outcome misclassification and selection bias on their findings. For instance, Wheeler et al. [12] believed asthma misclassification to be nondifferential, and they believed their findings were not spurious results caused by misclassification. They also mentioned that enrollees may have been more likely to develop asthma than non-enrollees. To assess potential self-selection bias, a quantitative assessment was conducted by excluding self-identified individuals (e.g., individuals enrolling through media or community outreach) from analyses and reanalyzing using only list-identified participants (e.g., individuals enrolling through employment records).

While these bias evaluations may be fairly typical for epidemiologic study reports, they do not quantitatively assess the combined effect of biases. Therefore, in this teaching article, we describe the data-generating process for WTCHR studies, and we apply this foundation to quantitatively [1,2] adjust a WTC study estimate for the combined impact of missing data, outcome misclassification, and nonparticipation. We use nonprobabilistic sensitivity analyses [1] to quantify the collective bias effects.

Data-generating process

We describe the components of the data-generating process, both in general and specifically for our WTCHR bias analysis.

1. Describe the research question (noting that the data-generating process may be different for different research questions).
   One research question Wheeler et al. [12] examined was as follows: What is the effect of exposure to the 9/11 dust cloud on self-reported newly physician-diagnosed asthma after 9/11?
2. Describe the target population. The target population is the population that investigators are interested in making inferences or asking questions about.
   The target population for Wheeler et al. [12] appears to be rescue-recovery workers and volunteers eligible for the Health Registry study of asthma.
3. Understand data collection and timing of human subject consent. For epidemiologic studies, data can be collected from multiple sources (e.g., laboratories, electronic health records, questionnaires, environmental data logs). Human subject consent may be obtained before, during, or after data collection.
   The WTCHR is a longitudinal study. Baseline interviews were completed using computer-assisted telephone and in-person interviewing, which included informed consent [10]. Wheeler et al.'s [12] analysis focused on the 2003–2004 baseline survey.
4. Understand errors in the data. List errors that occurred during the study design, data collection, and analyses.
   In this teaching article, we focus on the combined impact of selection bias (missing data and nonparticipation) and outcome misclassification.

Bias and fourfold tables

Using these data-generating components, we built upon the selection-bias research of Kleinbaum et al. [13] and expansion by Maldonado [14] and modified the selection-bias fourfold tables they presented. The selection-bias tables are placeholders for normal activities (i.e., selection forces) occurring in epidemiologic studies. We customized the selection-bias fourfold tables of Kleinbaum et al. and Maldonado to meet the needs of our bias analysis.

For the 9/11 dust cloud-asthma question of interest, the data-generating process simplified into three selection-bias fourfold tables (Fig. 1). "Target Population (Rescue-Recovery Workers & Volunteers Eligible for Health Registry Study of Asthma)" corresponds to rescue-recovery workers and volunteers who met the WTCHR asthma health study's eligibility criteria. "Participants (consented) in Health Registry Study of Asthma" are rescue-recovery workers and volunteers who consented, were eligible for the asthma study, and completed the WTCHR baseline interview. "2 × 2 Table for Data

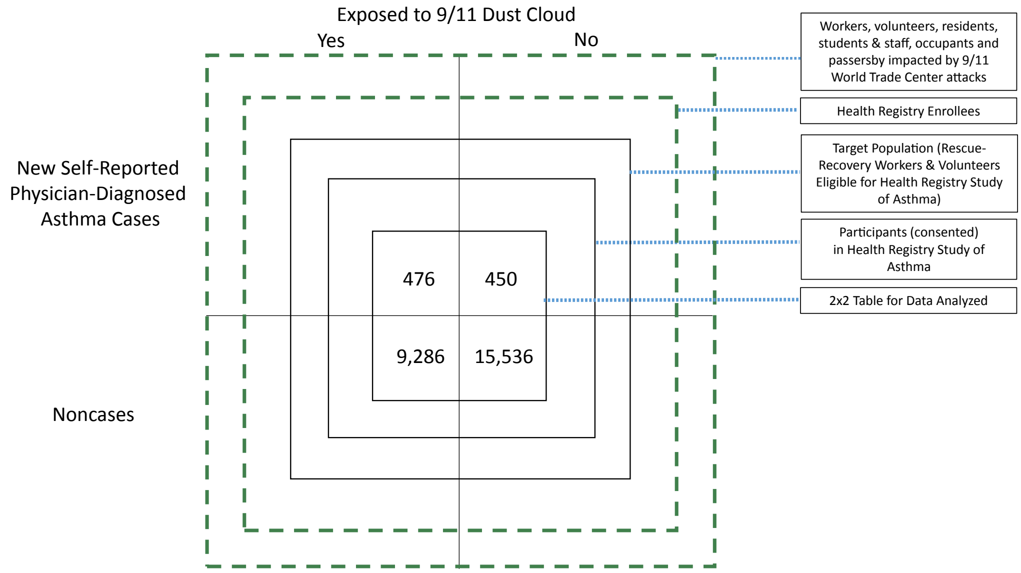

Fig. 1. Selection-bias fourfold tables. Numbers of cases from Wheeler et al. [12]. Numbers of noncases were not provided; estimated from given data.

Analyzed" are consented participants with complete data for the 9/11 dust cloud-asthma analysis.

For the WTCHR, neither the "Followed" nor "Sampled into Study" groups occurred (per the notation of Maldonado [14]). A contact list (i.e., "Followed") does not exist because the disaster was not expected. As a result, no predefined, identifiable source population was available to develop the cohort for surveillance. (There were after-the-fact efforts to identify WTCHR eligible participants based on employment or residential records [9].) Furthermore, no (random) sampling occurred because the voluntary [9] registry sought all eligible individuals. Therefore, we eliminated the "Followed" and "Sampled into Study" tables.

We included two larger dashed fourfold tables around the "Target Population (Rescue-Recovery Workers & Volunteers Eligible for Health Registry Study of Asthma)" table to indicate that there are additional people who belong outside the target population (Fig. 1). The inner dashed table ("Health Registry Enrollees") represents all participants in the Health Registry, whereas the outer dashed table ("Workers, volunteers, residents, students & staff, occupants and passersby impacted by 9/11 World Trade Center attacks") represents everyone effected by the terrorist attacks on 9/11.

The process so far assumed no errors in data. We therefore identified when misclassification could occur. In our example, classification error (for both the exposure and outcome) happened when the consented participants self-reported incorrectly while completing the survey, which transpired before any data were considered missing and the data were analyzed (Fig. 2). We indicated misclassification by replacing the "Participants (consented) in Health Registry Study of Asthma" table with a fourfold circle (Fig. 3). We then inserted a new fourfold table above the circle to account for the corrected data; we called this table "Correctly-Classified Participants."

Figure 3 illustrates the order that biases occur in the WTCHR, beginning with the outer 2 × 2 table and ending with the inner

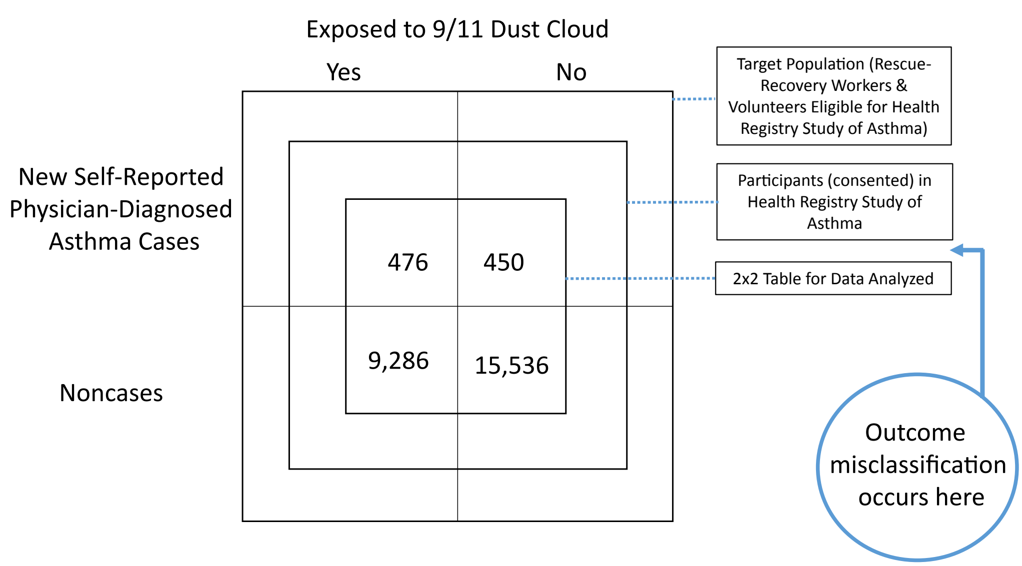

Fig. 2. Identifying misclassification during selection bias for a World Trade Center Health Registry Study of Asthma. Numbers of cases from Wheeler et al. [12]. Numbers of noncases were not provided; estimated from given data.

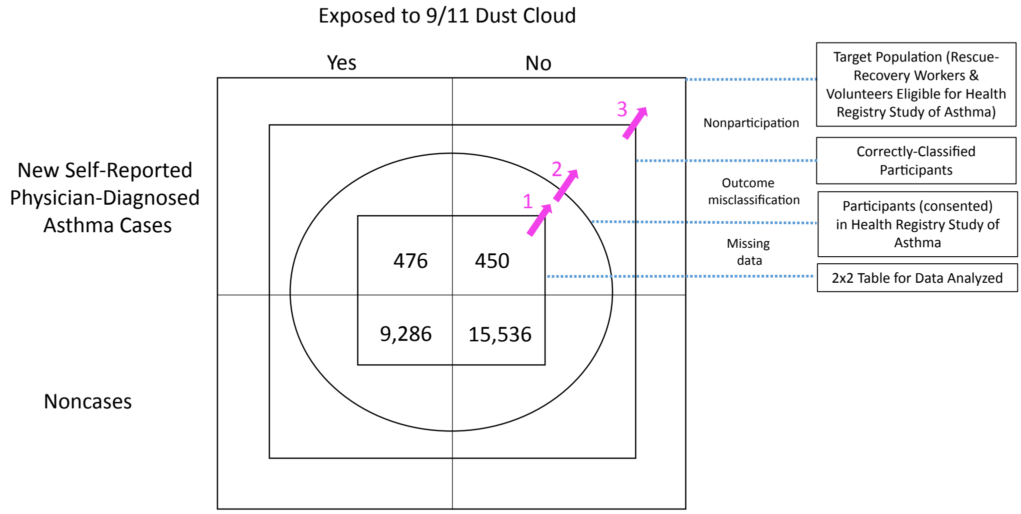

Fig. 3. Quantitative bias analysis adjustment steps for a World Trade Center Health Registry Study of Asthma. Numbers of cases from Wheeler et al. [12]. Numbers of noncases were not provided; estimated from given data.

2 × 2 table. That is, the first bias that occurs in the WTCHR is nonparticipation, followed by outcome misclassification, and then missing data. Therefore, when adjusting for these biases collectively in a bias analysis, the reverse sequence is used [1,2]. Additionally, we assume that once we adjusted for a bias, the data corrected for that bias are without that error.

Bias analysis

Wheeler et al.'s [12] 9/11 dust cloud-asthma 2 × 2 analysis comprised 25,748 people who had complete data on exposures, outcomes, and covariates. Their analysis did not attempt to correct for error in study results that may have been caused by selection bias or outcome misclassification. Therefore, their reported results are conditional on the implicit assumption that selection bias and outcome misclassification did not cause any important amount of error in the reported estimates. The goal of our bias analysis was to examine how sensitive Wheeler et al.'s results could have been to deviations from this implicit assumption.

To implement our quantitative bias analysis (i.e., sensitivity analysis), we found additional WTCHR data in the literature [8,10]. We specified various scenarios about the biases. We then quantitatively adjusted for bias in the reverse sequence of errors [1,2]. In Figure 3, we indicate this sequence by numbered arrows. Specifically, we started with the inner 2 × 2 table of data and worked outward adjusting first for missing data, followed by outcome misclassification and lastly nonparticipation.

Bias adjustment for missing data

Missing data refer to incomplete data. Participants, for instance, may skip certain questions when completing a questionnaire. Often when analyzing data, epidemiologists only use participants with complete data on all analyzed variables (as did Wheeler et al. [12]); yet certain variables may still be complete. Farfel et al. [10] reported enrollee characteristics, risk factors, exposures, and injuries for those completing the WTCHR baseline interview. There were 30,665 rescue-recovery workers and volunteers enrolled in the WTCHR. Of these, 30,525 answered the 9/11 dust cloud question (yes: $N = 11,355$; no: $N = 19,170$), whereas Wheeler et al. [12] analyzed fewer people (yes: $N = 9762$; no: $N = 15,986$). (Note we excluded from our bias analyses the 140 individuals (=30,665 − 30,525) who did not answer the 9/11 exposure question.)

Farfel et al. [10] also reported 971 reported cases of physician-diagnosed asthma among 30,546 rescue-recovery worker and volunteer enrollees. The number of cases in the Wheeler et al. [12] 2 × 2 table was 926 (= 476 + 450), 45 shy of the 971 reported by Farfel et al. [10] (indicating many more noncases had incomplete data). Our quantitative bias analysis for missing data involved making different assumptions about these 45 cases, which we assumed were individuals within the analyzed data set of 30,525. Because we did not know the exposure statuses, we chose the extreme scenarios: (a) all 45 are exposed cases and (b) all 45 are unexposed cases. Then to calculate the noncases, we subtracted the 45 cases and the observed cases from the number of rescue-recovery worker and volunteer enrollees who answered the 9/11 dust cloud question as reported by Farfel et al. [10]. For instance in scenario a, the adjusted number of noncases present during the 9/11 dust cloud (i.e., exposed noncases) was 10,834 (= 11,355 − 476 − 45). Once the fourfold tables were completed with the adjustment for missing data, we calculated the ORs adjusted for missing data, OR_M.

Bias adjustment for outcome misclassification

Self-reporting is prone to over- or under-reporting error. For our analyses, we sought classification probabilities (sensitivity and specificity) for self-reported physician-diagnosed asthma in the two exposure groups, exposed to 9/11 dust cloud ("yes" or "no"). We found no validation data for these groups. Because standard asthma questions were used in the WTCHR and asthma is a common medical term, we may expect high sensitivity values for self-reported physician-diagnosed asthma. However, the diversity of people in the dust cloud warrants consideration of response validity. For instance, the WTCHR contains workers (e.g., firefighters) who could lose their jobs because of health conditions resulting from disaster response. As a result, health conditions in certain professions could be under-reported to protect employment [15]. Conversely, volunteers and nonprofessionals may more likely over-report to ease their (future) livelihood fears due to unknown 9/11 exposures.

Our literature search found a validation study [16] from the New York City Fire Department whose firefighters were enrolled in the WTC Monitoring Program. The data compared self-reported physician-diagnosed asthma in 2005–2012 with medical records;

sensitivity was 68.7% (in both entire and restricted firefighter populations) and specificity values were 94.1% (entire population) and 93.2% (restricted population) [16]. We found two other validation studies [17,18] from European nondisaster populations self-reporting asthma diagnosis by health professionals with sensitivity and specificity more than 90%. Although professionals (e.g., firefighters and police) may have been more likely to have been present in the 9/11 dust cloud, less secure positions (e.g., transient workers and volunteers) may not have been. We chose one sensitivity value for those in the cloud (Se_{cloud} = 85%) and pivoted around it values in increments of 5%, with an emphasis on better recall in the noncloud participants ($Se_{noncloud}$) because less secure positions may report more accurately. Specificity values were limited (Sp_{cloud} and $Sp_{noncloud}$ = 1.0 and 0.98) due to the large number of noncases. That is, mathematically certain combinations of sensitivities, specificities, and observed cell counts used in bias analysis for correction of misclassification have potential for implausible (negative) values [1,19]. We used 16 outcome misclassification scenarios. For example, for Se_{cloud} = 0.85, $Se_{noncloud}$ = 0.95, Sp_{cloud} = 1.00, $Sp_{noncloud}$ = 0.98, the adjusted number of cases present during the 9/11 dust cloud (i.e., exposed cases) was 612.9412 (Fig. 4). We then calculated ORs adjusted for both missing data and outcome misclassification, OR_{MO}, using a matrix approach for misclassification adjustment [1,20].

Bias adjustment for nonparticipation

People eligible for health studies decide not to enroll for multiple reasons, including lack of interest or relocation. Wheeler et al. [12] analyzed 25,748 workers and volunteers, whereas Murphy et al. [8] estimated that 91,469 people helped with rescue and recovery efforts. We considered the latter number as the target population for rescue-recovery workers and volunteers. Using the fourfold table of correctly-classified participants, we calculated 60,944 nonparticipants (= 91,469 − 30,525). We found no literature data on nonparticipants' exposure and outcome statuses. (For instance, a method to obtain these data is by calling a sample of list-identified nonparticipants and requesting the information.) Accordingly, we were left with a large number of nonparticipants to accurately arrange (i.e., specify bias parameters for scenarios) in the fourfold table. However, we were not confident which scenarios could be plausible, and we had no information to indicate which scenarios would be more likely. Furthermore, we believed that certain fourfold-table arrangements of nonparticipants could drastically impact an OR estimate.

We thus identified one nonparticipation bias scenario for purposes of illustration. We assumed that most of the nonparticipants would be noncases and unexposed to the 9/11 dust cloud. Of the 60,944 nonparticipants, we specified 40,000 as unexposed noncases and 20,000 as exposed noncases. The remaining 944 were equally divided into exposed and unexposed cases. These fourfold-table counts were then added to the respective "Correctly-Classified Participants" counts before calculating the OR adjusted for missing data, outcome misclassification, and nonparticipation (OR_{MON}). All analyses were conducted using R statistical computing software (Vienna, Austria) [21].

Results

The nonprobabilistic sensitivity analyses results are found in Table 1. We calculated ORs adjusted for missing data (OR_M); missing data and outcome misclassification (OR_{MO}); and missing data,

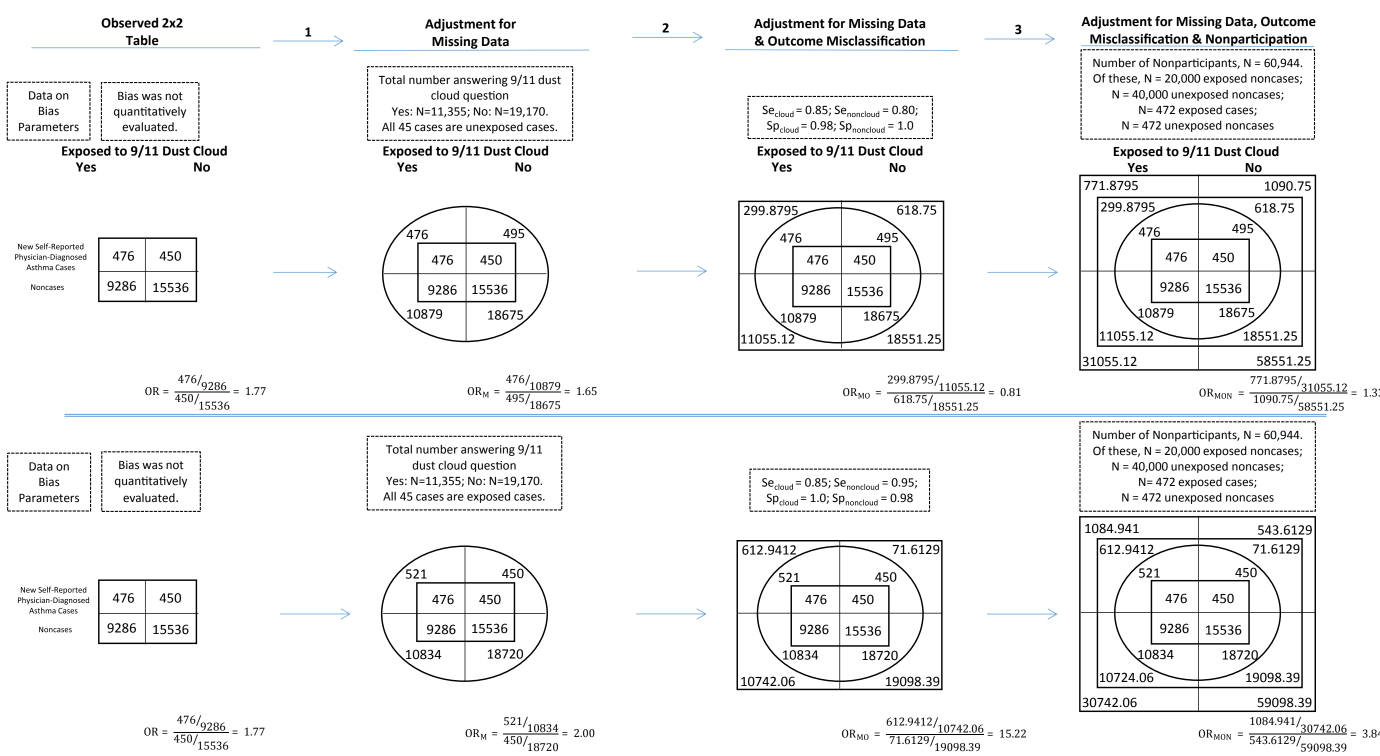

Fig. 4. Step-by-step examples of cell frequencies and odds ratios calculations for adjusting an odds ratio for missing data, outcome misclassification, and nonparticipation using data from the World Trade Center Health Registry. For the observed 2 × 2 table, numbers of cases are from Wheeler et al. [12]. Numbers of noncases were not provided; estimated from given data. OR_M = odds ratio adjusted for missing data; OR_{MO} = odds ratio adjusted for missing data and newly self-reported physician-diagnosed asthma (outcome) misclassification; OR_{MON} = odds ratio adjusted for missing data, newly self-reported physician-diagnosed asthma misclassification, and nonparticipation; Se_{cloud} = outcome sensitivity for rescue-recovery workers and volunteers self-reporting exposure to 9/11 dust cloud; $Se_{noncloud}$ = outcome sensitivity for rescue-recovery workers and volunteers self-reporting not exposed to 9/11 dust cloud; Sp_{cloud} = outcome specificity for rescue-recovery workers and volunteers self-reporting exposure to 9/11 dust cloud; $Sp_{noncloud}$ = outcome specificity for rescue-recovery workers and volunteers self-reporting not exposed to 9/11 dust cloud.

Table 1
Odds ratios adjusted for missing data, outcome (self-reported, newly diagnosed asthma) misclassification, and nonparticipation in a World Trade Center Health Registry Study of Asthma (unadjusted odds ratio = 1.77, 95% confidence limits (1.55, 2.01), as reported by Wheeler et al. [12])

			Sp_{cloud} = 1.0, $Sp_{noncloud}$ = 1.0		Sp_{cloud} = 1.0, $Sp_{noncloud}$ = 0.98		Sp_{cloud} = 0.98, $Sp_{noncloud}$ = 1.0		Sp_{cloud} = 0.98, $Sp_{noncloud}$ = 0.98	
	Se_{cloud}	$Se_{noncloud}$	OR_{MO}	OR_{MON}	OR_{MO}	OR_{MON}	OR_{MO}	OR_{MON}	OR_{MO}	OR_{MON}
Missing data scenario* a: OR_M = 2.00	0.85	0.95	2.25	2.19	15.22	3.84	1.27	1.65	8.58	2.90
	0.85	0.90	2.13	2.13	14.40	3.81	1.20	1.61	8.12	2.88
	0.85	0.85	2.01	2.07	13.57	3.78	1.13	1.56	7.66	2.85
	0.85	0.80	1.89	2.00	12.75	3.74	1.06	1.51	7.19	2.82
Missing data scenario b: OR_M = 1.65	0.85	0.95	1.86	1.98	8.24	3.34	0.97	1.47	4.31	2.48
	0.85	0.90	1.76	1.92	7.79	3.30	0.92	1.43	4.07	2.45
	0.85	0.85	1.66	1.86	7.34	3.26	0.87	1.38	3.84	2.42
	0.85	0.80	1.56	1.80	6.90	3.22	0.81	1.33	3.61	2.39

OR_M = odds ratio adjusted for missing data; OR_{MO} = odds ratio adjusted for missing data and newly self-reported physician-diagnosed asthma (outcome) misclassification; OR_{MON} = odds ratio adjusted for missing data, newly self-reported physician-diagnosed asthma misclassification, and nonparticipation; Se_{cloud} = outcome sensitivity for rescue-recovery workers and volunteers self-reporting exposure to 9/11 dust cloud; $Se_{noncloud}$ = outcome sensitivity for rescue-recovery workers and volunteers self-reporting not exposed to 9/11 dust cloud; Sp_{cloud} = outcome specificity for rescue-recovery workers and volunteers self-reporting exposure to 9/11 dust cloud; $Sp_{noncloud}$ = outcome specificity for rescue-recovery workers and volunteers self-reporting not exposed to 9/11 dust cloud.
* Missing data scenario a: 45 cases are exposed and scenario b: 45 cases are unexposed.

outcome misclassification, and nonparticipation (OR_{MON}). Under our assumptions about the selection bias and outcome misclassification bias parameters (and assuming no other errors), our ORs adjusted for all three biases (OR_{MON}) ranged from 1.33 to 3.84. Most of the adjusted estimates were greater than the observed value (OR = 1.77) and were outside the 95% confidence interval (1.55, 2.01). ORs adjusted for missing data and outcome misclassification produced a wider range of values, including protective effects and estimates greater than eight times the observed OR (OR_{MO} range: 0.81–15.21). Adjustment for missing data resulted in one estimate greater than the observed value (OR_M = 2.00), while the other was smaller (OR_M = 1.65). Figure 4 illustrates step-by-step bias adjustment of the cell frequencies and ORs calculations for two examples, OR_{MON} = 1.33 and OR_{MON} = 3.84.

Discussion

The 9/11 attacks on the World Trade Center and surrounding areas resulted in illnesses, injuries, and deaths among rescue-recovery workers and volunteers. For participants in the WTCHR, self-reporting pre- and post-9/11 health conditions and events may have resulted in inaccurate data and may possibly have caused bias. In this manuscript, we described the data-generating process for the WTCHR study and used that as the foundation for a quantitative bias analysis. We then experimented with bias-parameter values for missing data, outcome misclassification, and nonparticipation. Our results adjusted for all three biases collectively provided adjusted estimates that were generally larger than the observed OR value of 1.77. Without supporting data on bias parameters, however, the results are dependent on the validity of our assumptions about the bias parameters.

We lacked three key sets of information that would have been useful for bias analysis: complete data on participants, outcome (disease) validation data, and data on nonparticipants. Without these data to guide the quantitative bias analysis, we therefore speculated about, rather than specified with confidence, values for the bias parameters. Nonetheless, we still presented the bias analysis because understanding the unusual situations in man-made disaster epidemiology and thinking through the steps provides insight to design better epidemiologic studies from man-made disasters.

The WTCHR was created to monitor the health of those exposed to the WTC disaster. Registries such as the WTCHR provide valuable, descriptive reports for public health initiatives. Conversely, even when deemed for research rather than surveillance [9], registry data are not necessarily designed for estimating associations, generating hypotheses, or studying etiology. When researchers use registry data for such research, it requires careful assessment of the data and limitations.

Pre-existing environmental monitoring equipment may be destroyed during man-made disasters. Additionally, capturing environmental data soon after a disaster may be difficult. Therefore, registry data of man-made disaster exposures generally lack specificity and thus rely on proxies [22,23]. For instance, exposure to the 9/11 dust cloud as a "yes" or "no" response could have been interpreted numerous ways by those answering the WTCHR question. In fact, the WTCHR codebook cautions that participants may have had various interpretations of the dust cloud [24]. Exposure misclassification could be more than 50%. It therefore warrants a strong need for validation data when broad, unspecific questions are used, or additionally when questions having potential job implications (e.g., asthma diagnosis reported by firefighters) are solicited. Incorporating substudies [19] to assess data validity is recommended, and validation studies of particular health outcomes are being conducted [9]. Furthermore, given the diversity of people affected by disasters, finding applicable validation data in the literature may be difficult. One suggestion for analyses would be to separate participants by occupation. For example, instead of collapsing multiple professions into one broad category, such as rescue-recovery workers and volunteers, include additional analyses by job title.

Future disaster epidemiologic studies may also benefit from a proactive approach that focuses on the technical aspect of data collection. Assembling national and international public health scientists, including epidemiologists, to pre-identify (research) questions and data and to establish surveillance systems could initiative research before a catastrophe (either man-made or natural). As a result, any urgent need to create a registry or surveillance system after a disaster would be replaced with having an established foundation and supplementing it with additional surveillance or retrospective studies to collect current disaster-specific data. Creating data commonalities around all disasters would allow comparisons across disasters. Additionally, collaborating with emergency response teams (e.g., local medical reserves) and gathering information about emergency preparedness could benefit data collection efforts, particularly if existing infrastructure could be leveraged [25]. Moreover, forethought into using the data beyond the scope of surveillance or research for future planning and decision making should be considered. Funding and maintaining ongoing surveillance systems are critical for maintaining public health awareness.

One method to potentially cut costs is by augmenting registries with other sources of data. Syndromic surveillance systems actively monitor data in real time to identify potential signals of illnesses that may require further response or investigation [26,27]. One could build on syndromic surveillance by leveraging electronic sources of health care and environmental data and repurpose it for surveillance and clinical research. Interoperability is "the ability of health care information systems to work together and share information within and across organizational boundaries" [28, p. 44]. For instance, one could establish a secure transmission of consented patients' data from electronic health records to a registry. Interoperability can increase ascertainment, provide pre- and post-disaster data, monitor long-term follow-up, prevent duplication (and re-entry) of data that exist in other digital sources, and reduce self-reporting errors years after an event. Nonetheless, data from direct linkages may not be designed for research or be complete. Importantly, however, it could allow for data to be captured from multiple sources to prevent dependent errors. WTCHR collects both exposure and outcome data from the same source; adjustment simultaneously for exposure and outcome misclassification thus requires accounting for error dependency [29,30].

For our quantitative bias analysis of nonparticipation, we did not compare characteristics of participants with nonparticipants, which is typically done in epidemiologic reports [6,8]. Additionally, selected individuals may be removed from an analysis to check robustness in results, as Wheeler et al. [12] did by discarding self-identified participants. Rather, for selection bias, we were only interested in knowing the exposure and outcome statuses of those with incomplete data (i.e., missing data) and those who did not participate. This approach allows for evidence of a possible distortion of effect (i.e., bias).

Two common biases that we did not analyze are self-selection bias and the healthy-worker effect. WTCHR recruiting efforts included lists of employees and volunteers participating in the rescue-recovery efforts. Other participants enrolled through self-identification via media and community outreach [12]. Participation may have depended on exposure for those completely covered in the dust. Additionally, individuals knowing themselves to be ill (or injured) might have participated in the study more than non-ill individuals, making participation dependent on outcome. The target population was rescue-recovery workers and volunteers, some of whom spent many hours daily for months helping with rescue, recovery, and clean-up. They were able to work extended work shifts given their good health, but prolonged exposures could ultimately be hazardous to their health. As a result, the healthy-worker effect has a different meaning in worker populations participating in man-made disaster studies. That is, initially healthy workers may develop illnesses and injuries *because* of their ability to work.

Disaster epidemiology is not unique in having methodological challenges, yet the field is still in its infancy. Accordingly, efforts have been made to improve disaster studies [22,25]. The methodological challenge of evaluating biases occurs in every area of epidemiology. Although commonly ignored [31], epidemiologists have been encouraged to quantitatively assess biases in study results [1,32]. Importantly, we have been advised to use sensitivity analyses when analyzing WTC exposures and outcome [5]. Quantitative bias analysis shows the likely impact of evaluated sources of bias and hence affords improved interpretation of study estimates, especially those used for decision making (e.g., allocating resources for injuries and illnesses from disasters). This teaching article has demonstrated a bias analysis and has shown the need for gathering information on bias (parameters) to provide meaningful interpretations of WTC data.

Acknowledgment

The authors would like to thank those who reviewed the article, including the anonymous reviewers, Dr. Cathy Baase, Dr. Thomas John Bender, and Paul Bork; and Dr. Sherry Baron, Dr. Manuel Cifuentes, Dr. Sander Greenland, Dr. Hyun Kim, Dr. David Kriebel, Steve Mongin, and Andy Ryan for discussions on this topic. Also, the authors thank Dr. Matthew Fox, Dr. Aliza Fink, and Dr. Timothy Lash for developing bias analysis spreadsheets (https://sites.google.com/site/biasanalysis/) to check analyses produced in this work.

This work was supported by the National Institutes of Health—National Institute for Occupational Safety and Health to Dr. Hyun Kim, grant number U01OH010730.

References

[1] Greenland S, Lash TL. Bias analysis [Ch. 19]. In: Rothman K, Greenland S, Lash T, editors. Modern Epidemiology. 3rd ed. Philadelphia: Lippincott-Williams-Wilkins; 2008. p. 345–80.
[2] Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. Dordrecht: Springer; 2009.
[3] Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. Int J Epidemiol 2014;43(6):1969–85.
[4] Phillips C, Maldonado G. Using Monte Carlo methods to quantify the multiple sources of error in studies [Abstract]. Am J Epidemiol 1999;149:S17.
[5] Savitz DA, Oxman RT, Metzger KB, Wallenstein S, Stein D, Moline JM, et al. Epidemiologic research on man-made disasters: strategies and implications of cohort definition for World Trade Center worker and volunteer surveillance program. Mt Sinai J Med 2008;75(2):77–87.
[6] Yu S, Brackbill RM, Stellman SD, Ghuman S, Farfel MR. Evaluation of non-response bias in a cohort study of World Trade Center terrorist attack survivors. BMC Res Notes 2015;8:42.
[7] Brackbill RM, Thorpe LE, DiGrande L, Perrin M, Sapp JH, Wu D, et al. Surveillance for World Trade Center disaster health effects among survivors of collapsed and damaged buildings. MMWR Surveill Summ 2006;55(2):1–18.
[8] Murphy J, Brackbill RM, Thalji L, Dolan M, Pulliam P, Walker DJ. Measuring and maximizing coverage in the World Trade Center Health Registry. Stat Med 2007;26(8):1688–701.
[9] Cone JE, Farfel M. World Trade Center Health Registry—a model for a nanomaterials exposure registry. J Occup Environ Med 2011;53(6 Suppl):S48–51.
[10] Farfel M, DiGrande L, Brackbill R, Prann A, Cone J, Friedman S, et al. An overview of 9/11 experiences and respiratory and mental health conditions among World Trade Center Health Registry enrollees. J Urban Health 2008;85(6):880–909.
[11] The City of New York. 2015 News: Registry News Updates. 2015. Available at: http://www.nyc.gov/html/doh/wtc/html/news/news.shtml. Accessed March 9, 2016.
[12] Wheeler K, McKelvey W, Thorpe L, Perrin M, Cone J, Kass D, et al. Asthma diagnosed after 11 September 2001 among rescue and recovery workers: findings from the World Trade Center Health Registry. Environ Health Perspect 2007;115(11):1584–90.
[13] Kleinbaum DG, Morgenstern H, Kupper LL. Selection bias in epidemiologic studies. Am J Epidemiol 1981;113(4):452–63.
[14] Maldonado G. Adjusting a relative-risk estimate for study imperfections. J Epidemiol Community Health 2008;62(7):655–63.
[15] Gordon SB, Curran AD, Murphy J, Sillitoe C, Lee G, Wiley K, et al. Screening questionnaires for bakers' asthma—are they worth the effort? Occup Med (Lond) 1997;47(6):361–6.
[16] Weakley J, Webber MP, Ye F, Zeig-Owens R, Cohen HW, Hall CB, et al. Agreement between obstructive airways disease diagnoses from self-report questionnaires and medical records. Prev Med 2013;57(1):38–42.
[17] Oksanen T, Kivimäki M, Pentti J, Virtanen M, Klaukka T, Vahtera J. Self-report as an indicator of incident disease. Ann Epidemiol 2010;20(7):547–54.
[18] Pedotti R, Losappio L, Farinotti M, Preziosi D, Tramacere I, Stafylaraki C, et al. Accuracy of a questionnaire for identifying respiratory allergies in epidemiological studies. Rhinology 2015;53(1):49–53.
[19] Jurek AM, Lash TL, Maldonado G. Specifying exposure classification parameters for sensitivity analysis: family breast cancer history. Clin Epidemiol 2009;1:109–17.
[20] Greenland S, Kleinbaum DG. Correcting for misclassification in two-way tables and matched-pair studies. Int J Epidemiol 1983;12(1):93–7.
[21] R Development Core Team. R A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2015.
[22] Ekenga CC, Friedman-Jiménez G. Epidemiology of respiratory health outcomes among World Trade Center disaster workers: review of the literature 10 years after the September 11, 2001 terrorist attacks. Disaster Med Public Health Prep 2011;5 Suppl 2:S189–96.

[23] Moline JM, Herbert R, Levin S, Stein D, Luft BJ, Udasin IG, et al. WTC Medical Monitoring and Treatment Program: comprehensive health care response in aftermath of disaster. Mt Sinai J Med 2008;75(2):67–75.

[24] NYC Health. World Trade Center Health Registry Wave 1 Survey (2003-2004) Public Use Data Set Codebook. Available at: http://sites.bu.edu/tec/files/2016/01/wave1-codebook.pdf. Revised 9/12/2013. Accessed January 11, 2016.

[25] Dominici F, Levy JI, Louis TA. Methodological challenges and contributions in disaster epidemiology. Epidemiol Rev 2005;27:9–12.

[26] Borchardt SM, Ritger KA, Dworkin MS. Categorization, prioritization, and surveillance of potential bioterrorism agents. Infect Dis Clin North Am 2006;20(2):213–25.

[27] Kman NE, Bachmann DJ. Biosurveillance: a review and update. Adv Prev Med 2012;2012:301408.

[28] Glaser J. Interoperability: the key to breaking down information silos in health care. Healthc Financ Manage 2011;65(11):44–6, 48, 50.

[29] Kristensen P. Bias from nondifferential but dependent misclassification of exposure and outcome. Epidemiology 1992;3(3):210–5.

[30] Chavance M, Dellatolas G, Lellouch J. Correlated nondifferential misclassifications of disease and exposure: application to a cross-sectional study of the relation between handedness and immune disorders. Int J Epidemiol 1992;21(3):537–46.

[31] Jurek AM, Maldonado G, Greenland S, Church TR. Exposure-measurement error is frequently ignored when interpreting epidemiologic study results. Eur J Epidemiol 2006;21(12):871–6.

[32] Greenland S. Multiple-bias modeling for observational studies [with discussion]. J R Stat Soc Ser A 2005;168(2):267–306.