Assessing Nonresponse Bias in Farm Injury Surveillance Data

Cheryl L. Beseler[1,*], Risto H. Rautiainen[1]

[1] Department of Environmental, Agricultural, and Occupational Health, College of Public Health, University of Nebraska Medical Center, Omaha, Nebraska, USA.
[*] Correspondence: chbeseler@unmc.edu.

HIGHLIGHTS

- Despite a response rate of about 18%, there was scant evidence of nonresponse bias.
- Farm and ranch characteristics of operations responding to a mailed survey were similar to those not responding.
- Responders were significantly more likely to be married but only slightly older and more educated than nonresponders.
- Earlier survey responders reported more injuries and greater severity when injured compared to later responders.

ABSTRACT. *Nonresponse bias in a survey can result in misleading estimates of agricultural injuries and can misdirect prevention efforts aimed at reducing the burden of injuries on farmers. Responders (n = 2,977) and nonresponders (n = 13,849) were compared based on demographics and agricultural production characteristics to identify underrepresented subgroups. Injury characteristics were compared between early (n = 1,667) and late (n = 1,309) responders. Methods accounted for correlated data, sample size inflation of p-values, and assessment of meaningful differences. Few differences were identified between responders and nonresponders. Responders differed from nonresponders by state of residence, and responders were more likely to be married. Other characteristics (age, gender, education, farm size, crops grown, animals raised) were similar across groups. Early responders reported more injuries and more often sought medical care for an injury than late responders. The differences identified between responders and nonresponders were minimal and not likely to create bias. Differential reporting of injury and injury severity between early and late responders is worthy of further investigation.*

Keywords. Agriculture, Epidemiologic methods, Farm, Injury, Nonresponse bias, Occupational health, Ranch, Safety, Surveillance, Survey.

Capturing agricultural injuries in a surveillance system is a challenge when survey response rates have declined over the past four decades (Brick and Williams, 2013; Galea and Tracy, 2007; Kreuter, 2013; Morton et al., 2012). A census approach to collecting agricultural injury data is not feasible, making randomization in the collection

of farm data important to validity. However, randomization fails in the presence of nonresponse. Random sampling is the foundation of inferential statistics, ensuring that any important differences between the sample and the population are distributed non-differentially across participants. Nonresponse can create impactful differences between the sample and the population that the surveillance data is attempting to represent. However, a low response rate does not necessarily suggest the presence of nonresponse bias (Groves, 2006; Groves and Peytcheva, 2008; 't Mennetje et al., 2011). There are many reasons for response propensity, i.e., the likelihood that a survey recipient will respond, and they may or may not be associated with the survey variables. Whether bias exists depends on the correlation between the variables under study and the response propensity.

Agricultural injury and disease data are necessary for monitoring changes in the health and safety of farmers and farm workers. Assessment of the burden of injury and disease is the first step in burden, need, and impact (BNI), the evidence-based method of prioritizing research funds developed by the National Institute of Occupational Health and Safety (NOSH) (Felknor et al., 2019). The validity of injury surveillance data is uncertain because of a lack of reporting on the assessment of nonresponse bias for agricultural health and safety surveys.

The studies that have been published on nonresponse bias were primarily undertaken in the 1980s through the 2000s. An occupational study of nonresponders in a sample of 10,000 New Zealanders found that younger individuals responded at significantly lower rates than older individuals ('t Mennetje et al., 2011). Housewives, unemployed, retired, and indigenous people were underrepresented in the responders, which is not likely to create a bias in a study examining workplace exposures and diseases. After an initial mailing and a follow-up phone call (37% response rate), no significant differences in exposure (dust, oil/solvents, and pesticides) or health outcomes (asthma, sleep problems, lower back pain) were observed after comparing the crude prevalence rates based on actual survey responses to a weighted prevalence rate based on nonresponse. An examination of early and late responders showed no significant differences.

Socioeconomic status, demographic factors, and health status have been consistently observed to differ between responders and nonresponders (Heilbrun et al., 1991; Linnenkamp et al., 2020; Tolonen, et al., 2005). A 1976 study among nurses, with a response rate of 69.7%, found a 5% greater response rate among those who were younger (Barton et al., 1980). There was a slightly greater response rate among those with academic degrees, but no differences in employment status or state of residence. The authors speculated that the more homogeneous the target population, the lower the potential of response bias, but no formal studies have addressed this speculation.

Age is the most consistent difference in response propensity. A survey of 21,889 individuals designed to estimate the prevalence of disability in England found that response patterns differed by age group for those with a disability (Tennant and Bradley, 1991). In an initial mailing and three follow-up mailings (response rates of 57%, 73%, 81%, and 88%), disabled respondents ages 16 to 64 tended to respond sooner than disabled respondents over 64 years of age. The response for the over 64 group was relatively constant across follow-up requests. Reluctance to return a survey might be a function of both disease status and age, but regardless of age, additional follow-ups greatly increased participation.

In several studies of nonresponse bias of respiratory disease outcomes, younger male smokers had an increased probability of being nonresponders (Hazell et al., 2009; Kotaniemi et al., 2001; Ronmark et al., 2009). In studies of nonresponse of osteoporosis surveys

conducted in 1991 and 1995, biases were minimal. Nonresponders were older and less healthy than responders (Heilbrun et al., 1991). Early responders reported fewer risk factors for osteoporosis, more indicators of better health (less smoking, more exercise), and were more likely to have suffered a previous fracture than those responding later (O'Neill et al., 1995). The differences were less than 10% and did not alter the risk estimates for osteoporosis. However, those studies often assumed that late responders were similar to nonresponders when estimating effects of nonresponse. It is unclear whether this assumption is reasonable and whether it holds for an occupational health study of farmers.

The Central States Center for Agricultural Safety and Health (CS-CASH) conducted surveillance from 2011 to 2015 in collaboration with the USDA National Agricultural Statistics Service (NASS). Stratified random samples by state (Iowa, Kansas, Minnesota, Missouri, North Dakota, Nebraska, and South Dakota) were drawn by NASS from farm operations that responded to the most recent Census of Agriculture survey. The average response rate in the five annual surveys ($n = 34,777$ mailed out) was 32% (Johnson et al., 2021).

In 2018, CS-CASH updated and administered the Farm and Ranch Health and Safety Survey (FRHSS) in-house. A stratified random sample by state was then drawn by Farm Market iD (FMiD) from its database using two inclusion criteria: (1) the farm/ranch operation has an email address, and (2) the gross farm income (GFI) is at least $5,000. These criteria were applied to enable online surveys and to focus on active farm operations versus retired, discontinued, rural residence/acreage, land rented out, and other types of operations with low incomes from farming. The 2018 survey administered by CS-CASH demonstrated a lower response rate (19% vs. 32%) but a much higher injury rate (15% vs. 7%) compared to previous surveys administered by NASS, making it important to assess potential biases.

Including smaller farms is important to fully characterize farm injuries because 90% of U.S. farms are small family farms with less than $350,000 in gross cash farm income (Whitt et al., 2019). It is important to understand whether smaller farms are represented in the responders or whether there is a bias in favor of larger, better resourced farms because the risk of injury is likely to differ in important ways depending on farm characteristics. It is also important to assess demographic factors because personal characteristics such as age, gender, and education have been shown to be associated with injury (Dimich-Ward et al., 2004; Lizer and Petrea, 2008; Zwerling et al., 1995).

The smallest farms (<$5,000 total sales) represented 47.8% of the 2017 Ag Census operations (NASS, 2019) but only 19.1% of the respondents and 16.7% of the injury cases in the surveys administered by NASS during 2011 to 2015. The percentage of operations using internet access for day-to-day decision making was 76%, while the percentage of operations with an email address was not addressed in the 2017 Ag Census (NASS, 2019). A direct comparison between the Ag Census population and the FMiD database population after applying the two inclusion criteria cannot be made, but it could be assumed that approximately half of the farms in the 2018 FMiD data would not meet the two inclusion criteria. Due to low response in the <$5,000 sales category, the loss in potential responses in the 2018 survey is likely similar to the proportion of small farms that responded to the 2011-2015 surveys (19.1%).

This study was undertaken to better understand the quality of the surveillance data being collected. The primary objectives of this study were to (1) determine whether responders and nonresponders differed by demographic characteristics (state, gender, age, education, marital status, family size); (2) determine whether farm or ranch operator responders and nonresponders differed by farm type, size, or income level; and (3) determine whether

reported injuries among those who responded early differed from those who responded later, assuming that late responders are more similar to nonresponders than to early responders.

Materials and Methods

Study Sample

The sampling frame was obtained from Farm Market iD (FMiD, now DTN, https://www.dtn.com/agriculture/agribusiness/), a supplier of farm and grower data to agribusinesses, with over 2.4 million active agricultural operations in its database. The number of farm operations in the 2017 Census of Agriculture was 2.0 million (NASS, 2019). With differences in data collection and inclusion methods, FMiD's database is estimated to cover 95% of farms represented in the Ag Census in the CS-CASH surveillance region and nationally. A stratified random sample of at least 2,500 farms per state from Iowa, Kansas, Minnesota, Missouri, North Dakota, Nebraska, and South Dakota was requested from FMiD. The initial sampling frame included 22,440 farms in the CS-CASH region. As mentioned above, the inclusion criteria were having an email address and an estimated GFI of $5,000. The survey sample was reduced randomly to a maximum of 2,500 farm operations per state ($n = 16,826$), and these operations were contacted by letter in May 2018, asking them to complete and return the survey form (fig. 1) with a follow-up mailing to nonresponders in July 2018. Late responders were those who returned the survey on the second request, all of which were received after July 31, 2018. Data collection continued until November 2018. A total of 2,977 surveys were returned, resulting in a 17.7% response rate; 13,849 operations did not return a survey.

The surveillance research plan was reviewed by the University of Nebraska Medical Center institutional review board (IRB), which determined this surveillance research to be exempt (not human subjects research) (No. 452-11-EX).

Measures

Demographic Variables

Using the demographic data available from FMiD, we assessed differences in state of residence, gender, age, education, marital status, and number of children on the farm. Records in the FMiD database contain company names rather than individual names, resulting in missing data for some items. Gender was coded as female = 0 and male = 1; ambiguous responses and company names were coded as missing. Age was a continuous variable in the analysis. Education was coded as completed high school, completed college, completed
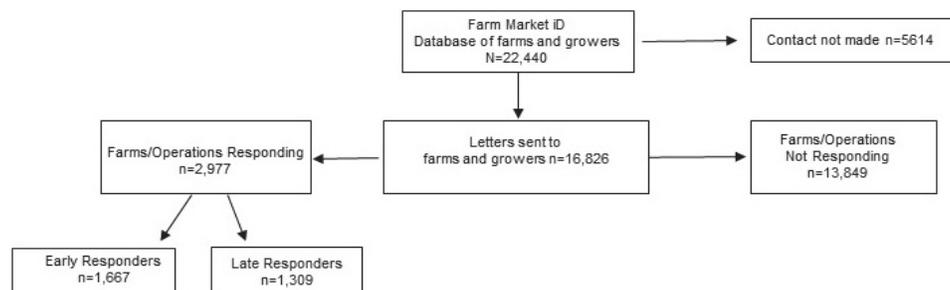

Figure 1. Survey sample drawn from Farm Market iD database in each response category, 2018.

graduate school, or attended a vocational or technical school. This variable was recoded into completed high school or attended vocational or technical school = 0 and completed college or graduate school = 1. Marital status was coded as single = 0 and married = 1, where single included inferred single, and married included inferred married. Number of children was a measure of family size where the operation was a family farm.

Farm Characteristic Variables

The 2018 FMiD data contained 56 farm production categories, some for single product types and some for combinations of products. For example, corn and soy were asked about separately and combined. The product responses were marked as either present or absent on the farm. To reduce the number of variables with sparse data, farm products were included if more than 30% of farm operations reported them as a farm product. Most commodities were reported as commonly produced or not at all. The result was 15 farm products retained for the nonresponse bias analysis. We used GFI as a measure of farm profitability. As a measure of farm size, we compared the number of planted acres and the number of acres harvested with a combine (combineable) between responders and nonresponders. We also compared the herd sizes of cattle between responders and nonresponders.

Farm Injury-Related Factors in Early and Late Responders

The FRHSS collects information for up to three operators per farm, including injury and illness outcomes, common agricultural exposures, and preventive measures. The unit of analysis for this study was the individual farmer identified in the survey as the person responding on behalf of the farm. We asked farm operators: "What percentage of this operator's work time was spent working on the farm/ranch in past 12 months?" with response categories of 0% to 24%, 25% to 49%, 50% to 74%, 75% to 99%, and 100%. Three injury and exposure questions were used to assess potential nonresponse bias: (1) How many farm injuries occurred to each operator in the past 12 months? (responses: 0, 1, 2, or 3 or more); (2) What professional medical care did this [most serious] injury require? (responses: none, doctor or clinic visit, hospitalization); and (3) When exposed to high levels of air contaminants at work, what percentage (%) of that time did the operator use proper respiratory protection (including N95 dust mask, half mask with dust/gas filters)? The date of July 31 was used to separate early responders from late responders. We conducted a sensitivity analysis and excluded 53 surveys received between July 24 and August 8 to assess whether larger differences between early and late responders altered our results.

Data Analysis

The many challenges of understanding nonresponse bias include assessing meaningful differences in responses because the p-values generated by statistics are more likely to be significant for large samples than for smaller samples (Good, 1980). Using p-values from traditional statistical tests will find differences due to the differing distributions of the p-value under the null and alternative hypotheses (Good, 1980, 1982; Mudge et al., 2012). There is also the problem of multiple testing of many correlated variables when making comparisons to assess differences between responders and nonresponders. In this study of nonresponse bias, we used principal component analysis (PCA) followed by a multivariate Kruskal-Wallis test to examine differences in farm production characteristics among responders and nonresponders. We used equivalence testing to assess meaningful differences in continuous farm and demographic variables.

We standardized the confidence level, alpha (α), to adjust for the large sample size of $n = 16,826$ in the responder/nonresponder analysis and $n = 4,333$ in the early/late responder analysis (Good, 1982). Therefore, the level of confidence used in these analyses for determining statistical significance was set at $\alpha = 0.004$ for the responder/nonresponder analysis and $\alpha = 0.01$ for the early/late responder analysis. With the exception of equivalence testing (Berger and Hsu, 1996), which was conducted in R (ver. 4.0.3, R Foundation), the demographic variables were analyzed using SAS (ver. 9.4, SAS Institute, Cary, N.C.). Multivariate normality was tested using the R package MVN (Korkmaz et al., 2014), and the farm characteristics comparisons were conducted in R using the micompr package (Fachada et al., 2016).

State of residence. We used the chi-square test to assess overall differences in state of residence of responders and nonresponders. Statistically significant differences were expected because CS-CASH has a greater presence in Nebraska. We calculated conditional probabilities to assess the likelihood of being a responder given residence in a certain state.

Gender, marital status, educational level, and number of children. Categorical demographic variables were tested using a chi-square test and an adjusted $\alpha = 0.004$.

Age. Age was a normally distributed variable in each response category, so a two-sample independent t-test was used to assess statistically significant differences. It was expected that differences would be identified, given the large sample size, so further analysis was conducted using equivalence testing (Berger and Hsu, 1996). The distribution in each of the groups was normally distributed, allowing two one-sided t-tests to be used to test whether the differences in ages were within five years. Five years was used because it is not a large enough spread to show differences in injury rates; most farm injury studies use wider ranges for age categories.

Farm income. Quantile-quantile (Q-Q) plots showed normality of the GFI variable after a natural log transformation. Values deviated from the Q-Q diagonal at very low farm incomes. Means and standard deviations (SD) are reported on the natural log scale, and the variances were considered to be unequal in a t-test.

Farm characteristics. We first checked for multivariate normality to assess whether a parametric or non-parametric method best fit the data distribution. The number of acres planted and the number of combineable acres were compared between responders and nonresponders using a two-sample non-parametric Wilcoxon approximation to the t-test due to being both positively skewed and kurtotic. With the goal of assessing farm characteristics as a cluster of correlated farm products where we could simultaneously compare a set of farm features among the two groups, we used PCA followed by a non-parametric Kruskal-Wallis multivariate analysis of variance (MANOVA) (Fachada et al., 2016; Langsrud, 2002). This method requires that the features be on the same scale. The crop variables were measured in number of acres and the animal variables were measured in herd size, requiring that these groups be analyzed separately. The first step in this approach is to replace the original set of variables with a few principal components that explain a desired amount of the variance in the sample. PCA converts multivariate responses into a set of linearly uncorrelated measures, with the algorithm selecting the features that best explain the sample differences. Where we had significant results in univariate Kruskal-Wallis tests, we calculated the eta-squared (η^2) statistics to assess effect sizes (Tomczak and Tomczak, 2014).

Farm Injuries. Differences in the number of operators on the farm, percentage of time each operator spent doing farm work, number of injuries on the farm in the past 12 months, the medical care required for the most serious injury, and the percentage of time using

personal protective equipment (PPE) when exposed to air contaminants were assessed using the chi-square test and Cochran-Armitage test for trends when a trend was suggested in ordinal variables. This analysis was done by operator and not by farm to capture injuries to all operators working on the farm and because the operator is the unit of analysis in published reports.

Results

Demographic Variables

Responders significantly differed from nonresponders by state of residence (p < 0.0001) (table 1). As expected, the highest response rates were in Nebraska and Iowa, where CS-CASH has a greater presence. The probability of responding was 0.27 if residing in Nebraska and 0.22 for Iowa; the remaining states ranged from probabilities of 0.13 to 0.17. Awareness of CS-CASH and its host, the University of Nebraska Medical Center, is the most likely explanation for these differences; response propensity was higher for those who were familiar with CS-CASH.

There were no statistically significant differences between responders and nonresponders for gender or number of children on the farm. However, responders were significantly more likely to be married than nonresponders (p = 0.0006) (table 1). The odds ratio for responding to the survey for those who were married compared to unmarried was 1.28 (99.6% confidence interval 1.04, 1.58), which is significant at α = 0.004. Responders were also more likely to have completed college or graduate school compared to nonresponders (p = 0.01), but this did not reach statistical significance.

Table 1. Demographic differences for responders (*n* = 2,977) and nonresponders (*n* = 13,849) in the survey sample extracted from the Farm Market iD database, 2018.[a]

Demographic Characteristics		Responders	Nonresponders
State of residence ***		(*n* = 2,977)	(*n* = 13,849)
	Iowa	550 (18.5)	1,953 (14.1)
	Kansas	352 (11.8)	2,126 (15.4)
	Minnesota	437 (14.7)	2,061 (14.9)
	Missouri	235 (7.9)	1,611 (11.6)
	North Dakota	334 (11.2)	2,166 (15.6)
	Nebraska	665 (22.3)	1,836 (13.3)
	South Dakota	404 (13.6)	2,096 (15.1)
Gender		(*n* = 2,906)	(*n* = 13,500)
	Female	167 (5.8)	847 (6.3)
	Male	2,739 (94.2)	12,653 (93.7)
Education level *		(*n* = 2,104)	(*n* = 9,897)
	High school/trade school	1,330 (63.2)	6,553 (66.2)
	College graduate or higher	774 (36.8)	3,344 (33.8)
Married **		(*n* = 2,270)	(*n* = 10,054)
	No	250 (11.0)	1,377 (13.7)
	Yes	2,020 (89.0)	8,677 (86.3)
Number of children		(*n* = 1,382)	(*n* = 6,607)
	1	834 (60.3)	3,809 (57.6)
	2	238 (17.2)	1,242 (18.8)
	3	158 (11.4)	808 (12.2)
	4	148 (10.7)	741 (11.2)
	5	4 (0.3)	7 (0.1)

[a] Values are means (percentages in parentheses). Asterisks indicate statistically significant differences between responders and nonresponders: * (p < 0.05), ** (p < 0.001), and *** (p < 0.0001).

Responders were significantly older than nonresponders (t = -6.42, p < 0.0001) despite the differences between the groups being slight. Equivalence testing showed that the average age difference between nonresponders (*n* = 10,522, mean = 59.35, SD = 12.4) and responders (*n* = 2,352, mean = 60.97, SD = 10.7) was -1.62 with a 90% confidence interval for a one-sided test of (-2.03, -1.20) and was clearly within the 2.5 equivalence bound allowed for a one-tailed test. The small age difference between responders and nonresponders is not likely to result in bias in assessing injury risk.

Farm Characteristics

GFI did not differ significantly between nonresponders (mean = 12.8, SD = 0.96) and responders (mean = 12.8, SD = 0.85) (t = -0.66, p = 0.51). The median income was $372,250 for nonresponders and $375,411 for responders. Neither the number of planted acres (t = -2.64, p = 0.008) nor the number of combineable acres (t = -2.37, p = 0.02) significantly differed between responders and nonresponders.

Two principal components explained nearly all of the variance in crops grown on the farms, indicating the homogeneous nature of commodities grown in the region. The first principal component explained 78.7% of the variance, and the second explained 13.5%, leaving 8% explained by small amounts across the remaining principal components. The component that explained the most variance did not differ between responders and nonresponders (p = 0.67), but the second and third components showed statistically significant differences. Due to more than 20% of the variance differing by response status, we examined each crop individually in a non-parametric Kruskal-Wallis test, calculated an effect size, and plotted the median values (fig. 2).

Of the 11 crops, only field corn, corn, and wheat met the criteria of having a p-value less than 0.004. The difference in acreage between corn and field corn was less than 25 acres. Wheat is not shown in figure 2 because the median was zero; 75% of the sample had fewer than 100 acres of wheat. The η^2 estimates for the effect sizes were all much less than 0.01, indicating a small effect size that was not likely to result in nonresponse bias.
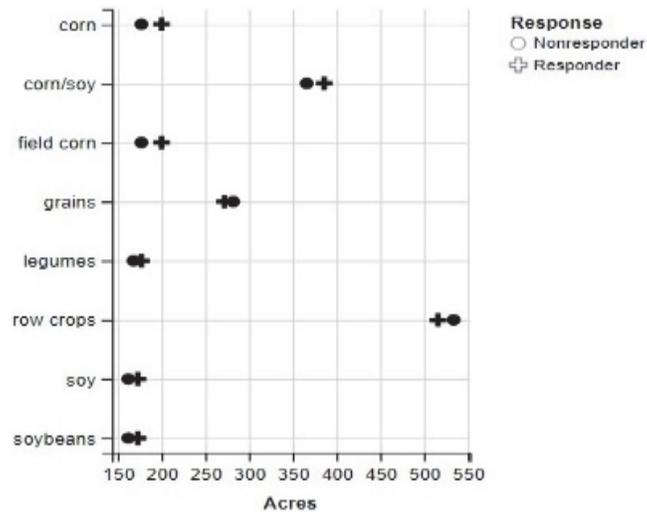

Figure 2. Median number of acres for farm crop products for responders and nonresponders to survey request after an initial mailing and a follow-up mailing, 2018.

Table 2. Operator and injury characteristics from early ($n = 2,473$) and late responders ($n = 1,860$), 2018.[a]

Operator and Injury Characteristics		Early Responders	Late Responders
Operators on the farm	1	1,667 (67.4)	1,309 (70.4)
	2	620 (25.1)	411 (28.1)
	3	186 (7.5)	140 (7.5)
Percentage of work time spent doing farm work	0% to 24%	1,278 (52.4)	992 (54.5)
	25% to 49%	483 (19.8)	359 (19.7)
	50% to 74%	259 (10.6)	181 (9.95)
	75% to 99%	276 (11.3)	182 (10.0)
	100%	144 (5.90)	105 (5.77)
Total number of injuries reported *	None	2,125 (86.2)	1,641 (89.2)
	1	276 (11.2)	167 (9.08)
	2	47 (1.91)	24 (1.31)
	3 or more	17 (0.69)	7 (0.38)
Medical care required **	None	2,146 (86.8)	1,667 (89.6)
	Doctor or clinic visit	327 (13.2)	193 (10.4)

[a] Values are means (percentages in parentheses). Asterisks indicate statistically significant differences between responders and nonresponders: * ($p < 0.05$), ** ($p < 0.01$).

A single factor explained 93.5% of the variance in the four animal production variables, with the second factor explaining 5.33%. The p-value for the association between being a responder and nonresponder was 0.76 for the first principal component and 1.00 for the second principal component. There were no differences in the medians of animals produced between responders and nonresponders in animal production.

Injury Variables

A total of 1,667 (56%) farms responded to the first request, and 1,309 (44%) responded to the second request. The return date on one survey was undetermined and was excluded from the analysis. No significant differences were observed between early and late responders in the number of primary, second, and third operators working on the farm (table 2). There were no statistically significant differences between early and late responders in number of reported injuries or percentage of work time spent doing farm work (table 2).

However, the trend test for the number of injuries suggested that early responders reported a greater number of injuries than late responders ($p = 0.002$). The median percentage of time wearing PPE when exposed to air contaminants was 50% for early and late responders and ranged from 0% to 100% for both groups ($\chi^2 = 0.02$, $p = 0.89$). Early responders were more likely to have visited a doctor or clinic due to an injury than late responders ($p = 0.004$) (table 2). For those who reported an injury, 13.2% of early responders reported visiting a doctor after an injury, compared to 10.4% of late responders. In the sensitivity analysis, excluding 53 who responded one week before or after the July 31 cutoff date (after July 24 and before August 8), the results remained unchanged.

Discussion

Few differences were identified between those who responded to the farm safety and health survey and those who did not. Responders were significantly more likely to be married, more frequently from Nebraska or Iowa, and slightly younger, and with greater education. No differences were observed by gender, family size, farm income, or farm size. Only small differences were seen in the types of farm crops produced and are not likely to be of sufficient magnitude to cause a bias in effect sizes in the injury and health outcomes of interest. Age has not been consistently associated with response propensity and showed

only a slight difference in this study. Our results are similar to the occupational study of 10,000 New Zealand workers, in which younger respondents were less likely to respond than older respondents ('t Mennetje et al., 2011).

Consistent with previous studies, age and education may be demographic differences in response propensity but may differ by population and outcome under study. Those who responded earlier were more likely to have had a serious injury than those who responded later. The numbers of injuries reported were similar among the two groups using a cutoff date of July 31, and this remained unchanged in the sensitivity analysis. In the trend analyses, early responders reported a significantly greater number of injuries than late responders. This is suggestive of the osteoporosis study comparing early responders to later responders, which found more previous fractures in the early responders (O'Neill et al., 1995).

Marital status has not been consistently reported as a risk factor for agricultural injury (Jadhav et al., 2016), and further clarification of marital status as a risk factor is needed. It is unclear why more married individuals responded when gender did not differ between responders and nonresponders. Having two people reminding one another to return the survey may increase the probability of responding. The 3% difference in being married or not married and responding to the survey is not likely to create a bias, but it is worth noting that unmarried farmers are likely to be underrepresented in the sample.

The homogeneity of the farming population in our seven-state region may be an explanation for the lack of evidence of nonresponse bias. As originally observed and tested by Wallace (1947), and as speculated by Barton et al. (1980) in their 1976 study of nonresponse among nurses, homogeneity of the sample may reduce bias in mailed questionnaires. Farmers in the midwestern U.S. are likely to be a homogeneous group. The reasons farmers may not return surveys are probably unrelated to the survey items and more likely reflect a dislike of surveys in general. Farmers experience time pressures and are independent thinkers. Many may feel as if the survey is intrusive, and they may not feel an obligation to divulge personal information about their farm operations. It is likely personality and psychological factors that result in whether a farmer returns a survey. It may also be the case that those who returned the survey have experienced a work-related injury or health issue and have an interest in sharing their story. They may have a greater concern for health and safety. It is unclear whether the late responders with fewer serious injuries are more similar to nonresponders, and there is no easy way to answer this question.

This study used several approaches to working with large samples and correlated data. Even after adjusting the p-value for the different distributions under the null and alternative hypotheses, the differences identified were not of large magnitude. It is not understood how different some of these farm characteristics need to be between responders and nonresponders to make a difference in risk estimates of injury and health outcome data. Determining what is a meaningful difference would require a more thorough investigation and would be a challenging undertaking. The 50-50 MANOVA approach simplified the analysis by providing data reduction for a number of highly correlated crop and animal variables. About 78% of the variance in crop production was accounted for in a single dimension, demonstrating the similarity of farming operations in the region. The remaining differences were significant but did not appear large enough to have an effect on injury risk estimation. These methods work well for large, correlated samples.

Equivalence testing is an approach that is useful but underused. Testing for whether a mean difference is equal to zero is not a useful exercise because differences will never be

zero and will be statistically significant for large samples. Testing whether a value falls within a predefined range is useful when testing for meaningful differences.

There are inherent limitations to this type of study. It is not possible to address the question of whether injuries are different between responders and nonresponders, and we have no information on whether nonresponders are similar to late responders. Operators with a recent injury might have been more likely to recall the injury when responding to the survey, or operators who experienced any injury might have a higher response propensity than those who did not experience an injury. In this study, 44% of the responders were late responders, so it is a leap to assume that this 44% was far different from the 56% who responded earlier. It is likely that the nonresponders differ in ways that cannot be adequately measured. Using a single time point as the cut-off date is not ideal because some of the surveys differed by only a day or two around the cut-off date. A sensitivity analysis without observations close to the cut-off date is helpful in examining the robustness of the early versus late responder comparison.

Farms were excluded in the initial sampling plan. The requirement of having an email address and at least $5,000 in GFI resulted in some farms being excluded, but it also prevented residential acreage and small farms with no active production from being included in the sample. The FMiD sampling frame contains 95% of the USDA Ag Census farms, and FMiD covers a relatively large number of farms that are not counted in the Ag Census. However, there is a time lag in updating large databases to reflect the changing farming environment, including sales of farmland and changes in production. These types of changes may result in the survey not reaching its intended target and may account for a small number of the nonresponses.

A possible solution to the state differences is oversampling in states other than Nebraska and Iowa. Other problems are more difficult to resolve. For this study, if the late responders resemble the nonresponders, there is evidence that serious injuries are overrepresented in the surveillance system. Injured operators who sought medical care were more likely to respond than those who did not seek care, but we do not know whether it was injury severity or the need for medical care that resulted in earlier responding. In addition, there is always the concern that individuals do not respond due to poor health or having had a serious injury related to farm work. This makes supplementing the surveillance data with other sources of records important to assess what might be missing (Earle-Richardson et al., 2011; Kica and Rosenman, 2020).

The future directions for documenting the validity and representativeness of the FRHSS will involve obtaining data from hospital discharge records, emergency room visits, and emergency medical transport services. Fatal injuries in agriculture, as in other industries, are enumerated by the Bureau of Labor Statistics in the Census of Fatal Occupational Injuries (CFOI). The CS-CASH surveillance program includes media monitoring of fatal and serious injury incidents reported in the media. This system identifies similar numbers of fatality cases as the CFOI, so there is some certainty that fatal injuries are well documented in agriculture in the CS-CASH region.

Conclusions

We found no evidence that nonresponse bias was compromising the quality of the prevalence estimates calculated from the surveillance system for agricultural injuries. With response rates in surveys continuing to fall and greater incentives becoming necessary to

solic responses, it is important for the integrity of these estimates to understand who might not be represented in the surveillance system data. Future research should assess whether late responders are more similar to nonresponders, or whether they are a group with their own set of response characteristics.

References

Barton, J., Bain, C., Hennekens, C. H., Rosner, B., Belanger, C., Roth, A., & Speizer, F. E. (1980). Characteristics of respondents and non-respondents to a mailed questionnaire. *American J. Public Health, 70*(8), 823-825. https://doi.org/10.2105/ajph.70.8.823

Berger, R. L., & Hsu, J. C. (1996). Bioequivalence trials, intersection-union tests, and equivalence confidence sets. *Stat. Sci., 11*(4), 283-319. https://doi.org/10.1214/ss/1032280304

Brick, J. M., & Williams, D. (2013). Explaining rising nonresponse rates in cross-sectional surveys. *Ann. American Acad. Polit. So. Sci., 645*(1), 36-59. https://doi.org/10.1177/0002716212456834

Dimich-Ward, H., Guernsey, J. R., Pickett, W., Rennie, D., Hartling, L., & Brison, R. J. (2004). Gender differences in the occurrence of farm-related injuries. *Occup. Environ. Med., 61*(1), 52-56.

Earle-Richardson, G. B., Jenkins, P. L., Scott, E. E., & May, J. J. (2011). Improving agricultural injury surveillance: A comparison of incidence and type of injury event among three data sources. *American J. Ind. Med., 54*(8), 586-596. https://doi.org/10.1002/ajim.20960

Fachada, N., Rodrigues, J., Lopes, V. V., Martins, R. C., & Rosa, A. C. (2016). micompr: An R package for multivariate independent comparison of observations. *The R Journal, 8*(2), 405-420. https://doi.org/10.32614/RJ-2016-055

Felknor, S. A., Schulte, P. A., Schnorr, T. M., Pana-Cryan, R., & Howard, J. (2019). Burden, need, and impact: An evidence-based method to identify worker safety and health research priorities. *Ann. Work Expos. Health, 63*(4), 375-385. https://doi.org/10.1093/annweh/wxz011

Galea, S., & Tracy, M. (2007). Participation rates in epidemiologic studies. *Ann. Epidemiol., 17*(9), 643-653. https://doi.org/10.1016/j.annepidem.2007.03.013

Good, I. J. (1980). C73. The diminishing significance of a p-value as the sample size increases. *J. Stat. Comput. Simul., 11*(3-4), 307-313. https://doi.org/10.1080/00949658008810416

Good, I. J. (1982). C140. Standardized tail-area prosabilities. *J. Stat. Comput. Simul., 16*(1), 65-66. https://doi.org/10.1080/00949658208810607

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opin. Qtly., 70*(5), 646-675. https://doi.org/10.1093/poq/nfl033

Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opin. Qtly., 72*(2), 167-189. https://doi.org/10.1093/poq/nfn011

Hazell, M., Morris, J., Linehan, M., Frank, P., & Frank, T. (2009). Factors influencing the response to postal questionnaire surveys about respiratory symptoms. *Primary Care Respir. J., 18*(3), 165-170. https://doi.org/10.3132/pcrj.2009.00001

Heilbrun, L. K., Ross, P. D., Wasnich, R. D., Yano, K., & Vogel, J. M. (1991). Characteristics of respondents and nonrespondents in a prospective study of osteoporosis. *J. Clin. Epidemiol., 44*(3), 233-239. https://doi.org/10.1016/0895-4356(91)90034-7

Jadhav, R., Achutan, C., Haynatzki, G., Rajaram, S., & Rautiainen, R. (2016). Review and meta-analysis of emerging risk factors for agricultural injury. *J. Agromed., 21*(3), 284-297. https://doi.org/10.1080/1059924X.2016.1179611

Johnson, A., Baccaglini, L., Haynatzki, G. R., Achutan, C., Loomis, D., & Rautiainen, R. H. (2021). Agricultural injuries among farmers and ranchers in the central United States during 2011-2015. *J. Agromed., 26*(1), 62-72. https://doi.org/10.1080/1059924X.2020.1845268

Kica, J., & Rosenman, K. D. (2020). Multisource surveillance for non-fatal work-related agricultural injuries. *J. Agromed., 25*(1), 86-95. https://doi.org/10.1080/1059924X.2019.1606746

Korkmaz, S., Goksolok, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *The R Journal, 6*(2), 151-162. https://doi.org/10.32614/RJ-2014-031

Kotaniemi, J.-T., Hassi, J., Kataja, M., Jonsson, E., Laitinen, L. A., Sovijarvi, A. R., & Lundback, B. (2001). Does non-responder bias have a significant effect on the results in a postal questionnaire study? *European J. Epidemiol., 17*(9), 809-817. https://doi.org/10.1023/A:1015615130459

Kreuter, F. (2013). Facing the nonresponse challenge. *Ann. American Acad. Polit. Social Sci., 645*(1), 23-35. https://doi.org/10.1177/0002716212456815

Langsrud, O. (2002). 50-50 multivariate analysis of variance for collinear responses. *J. Royal Stat. Soc. Series D, 51*(3), 305-317. https://doi.org/10.1111/1467-9884.00320

Linnenkamp, U., Gontscharuk, V., Brune, M., Chernyak, N., Kvitkina, T., Arend, W.,... Icks, A. (2020). Using statutory health insurance data to evaluate non-response in a cross-sectional study on depression among patients with diabetes in Germany. *Intl. J. Epidemiol., 49*(2), 629-637. https://doi.org/10.1093/ije/dyz278

Lizer, S. K., & Petrea, R. E. (2008). Health and safety needs of older farmers: Part II. Agricultural injuries. *AAOHN J., 56*(1), 9-14. https://doi.org/10.1177/216507990805600102

Morton, S. M. B., Bandara, D. K., Robinson, E. M., & Carr, P. E. A. (2012). In the 21st century, what is an acceptable response rate? *Australian New Zealand J. Public Health, 36*(2), 106-108. https://doi.org/10.1111/j.1753-6405.2012.00854.x

Mudge, J. F., Baker, L. F., Edge, C. B., & Houlahan, J. E. (2012). Setting an optimal α that minimizes errors in null hypothesis significance tests. *PLoS One, 7*(2), e32734. https://doi.org/10.1371/journal.pone.0032734

NASS. (2019). 2017 Census of Agriculture. Washington, DC: USDA National Agricultural Statistics Service. Retrieved from https://www.nass.usda.gov/AgCensus/

O'Neill, T. W., Marsden, D., & Silman, A. J. (1995). Differences in the characteristics of responders and non-responders in a prevalence survey of vertebral osteoporosis. *Osteoporos. Intl., 5*(5), 327-334. https://doi.org/10.1007/BF01622254

Ronmark, E. P., Ekerljung, L., Lotvall, J., Toren, K., Ronmark, E., & Lundback, B. (2009). Large-scale questionnaire survey on respiratory health in Sweden: Effects of late- and non-response. *Respir. Med., 103*(12), 1807-1815. https://doi.org/10.1016/j.rmed.2009.07.014

't Mannetje, A., Eng, A., Douwes, J., Ellison-Loschmann, L., McLean, D., & Pearce, N. (2011). Determinants of non-response in an occupational exposure and health survey in New Zealand. *Australian New Zealand J. Public Health, 35*(3), 256-263. https://doi.org/10.1111/j.1753-6405.2011.00703.x

Tennant, A., & Badley, E. M. (1991). Investigating non-response bias in a survey of disablement in the community: Implications for survey methodology. *J. Epidemiol. Commun. Health, 45*(3), 247-250. https://doi.org/10.1136/jech.45.3.247

Tolonen, H., Dobson, A., & Kulathinal, S. (2005). Effect on trend estimates of the difference between survey respondents and non-respondents: Results from 27 populations in the WHO MONICA Project. *European J. Epidemiol., 20*(11), 887-898. https://doi.org/10.1007/s10654-005-2672-5

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends Sport Sci., 1*(21), 19-25.

Wallace, D. (1947). Mail questionnaires can produce good samples of homogeneous groups. *J. Marketing, 12*(1), 53-60. https://doi.org/10.1177/002224294701200106

Whitt, C., MacDonald, J. M., & Todd, J. E. (2019). America's diverse family farms: 2019 edition. EIB-214. Washington, DC: USDA Economic Research Service. Retrieved from https://www.ers.usda.gov/publications/pub-details/?pubid=95546

Zwerling, C., Sprince, N. L., Wallace, R. B., Davis, C. S., Whitten, P. S., & Heeringa, S. G. (1995). Occupational injuries among agricultural workers 51 to 61 years old: A national study. *J. Agirc, Saf. Health, 1*(4), 273-281. https://doi.org/10.13031/2013.19469