

**Appendix D: Sampling Estimators Used for the Racial
Minority Farm Operator Childhood Agricultural Injury
Survey**

- Equations to derive the non-benchmarked estimates, y , and variances, $v(y)$, for injury, asthma, or youth for racial minority farms.

$$y = \sum_{g=1}^9 \sum_{h=1}^4 \left[\sum_{i=1}^{n_{gh}} \frac{N_{gh}}{n_{gh}} y_{ghi} \right] = \sum_{g=1}^9 \sum_{h=1}^4 y_{gh} \quad (\text{D.1})$$

$$v(y) = \sum_{g=1}^9 \sum_{h=1}^4 \left[\frac{N_{gh} - n_{gh}}{N_{gh}} \sum_{i=1}^{n_{gh}} \frac{(N_{gh} y_{ghi} - y_{gh})^2}{n_{gh}(n_{gh} - 1)} \right] = \sum_{g=1}^9 \sum_{h=1}^4 v(y_{gh}) \quad (\text{D.2})$$

where:

g = regional strata used in post-stratification;

h = racial strata used in post stratification;

N_{gh} = number of farms of race h in region g from the NASS sampling list;

n_{gh} = number of respondents of race h in region g from the survey;

y_{ghi} = value of the variable of interest (i.e., injury, asthma, youth) on farm i of race h in region g from the survey;

y_{gh} = estimate of variable of interest (i.e., injury, asthma, youth) for race h in region g ;

$v(y_{gh})$ = sampling variance for the variable of interest (i.e., injury, asthma, youth) for race h in region g .

- Equations to derive the benchmarked national estimates, $y_{(bm)}$, and variances, $v(y_{(bm)})$, for injury, asthma, or youth for racial minority farms.

$$y_{(bm)} = \sum_{g=1}^9 \sum_{h=1}^4 \left[\frac{N_{(bm)gh}}{N_{gh}} y_{gh} \right] \quad (\text{D.3})$$

$$v(y_{(bm)}) = \sum_{g=1}^9 \sum_{h=1}^4 \left[\left(\frac{N_{(bm)gh}}{N_{gh}} \right)^2 v(y_{gh}) \right] \quad (\text{D.4})$$

where non-benchmarked values are as previously defined in D.1 and D.2, and:

$N_{(bm)gh}$ = number of farms of race h in region g from the published 1997 Census of Agriculture.

Note: Summing equations D.3 and D.4 over the four racial strata, h , within a specific region, g , provides the benchmarked estimate for the variable of interest and its corresponding variance for all racial minority farms in region g .

Summing equations D.3 and D.4 over the nine geographic strata, g , for a specific racial group, h , provides the benchmarked national estimate for the variable of interest and its corresponding variance for racial group h .

3. Equations to derive the benchmarked national injury incidence or asthma prevalence rate estimates, R , and variances, $v(R)$, for racial minority farms.

$$R = 1000 \left(\frac{y_{(bm)}}{x_{(bm)}} \right) \quad (D.5)$$

$$v(R) = \frac{1000^2}{n} \left(\frac{1}{x_{(bm)}} \right)^2 \left[v(y_{(bm)}) + R^2 v(x_{(bm)}) - 2R \text{cov}(y_{(bm)}, x_{(bm)}) \right] \quad (D.6)$$

where:

$y_{(bm)}$ = benchmarked national estimate for injury or asthma from the survey;

$x_{(bm)}$ = benchmarked national estimate for youth at risk from the survey;

$\bar{x}_{(bm)}$ = benchmarked national average of youth at risk per farm from the survey;

$v(y_{(bm)})$ = benchmarked variance for the national injury or asthma estimate from the survey;

$v(x_{(bm)})$ = benchmarked variance for the national estimate of youth at risk from the survey;

$\text{cov}(y_{(bm)}, x_{(bm)})$ = covariance between the benchmarked injury or asthma estimate and the benchmarked estimate of youth at risk from the survey;

n = the number of farms from the NASS sampling frame used to derive the covariance between $y_{(bm)}$ and $x_{(bm)}$.

Alternatively, the variance for R can be determined by:

$$v(R) = 1000^2 R^2 \left[\left(\frac{\sqrt{v(y_{(bm)})}}{y_{(bm)}} \right)^2 + \left(\frac{\sqrt{v(x_{(bm)})}}{x_{(bm)}} \right)^2 - 2 \left(\frac{\text{cov}(y_{(bm)}, x_{(bm)})}{y_{(bm)} \bar{x}_{(bm)} n} \right) \right] \quad (D.6.1)$$

where:

$$\frac{\sqrt{v(y_{(bm)})}}{y_{(bm)}} = \text{relative standard error for } y_{(bm)} \text{ (i.e., injury or asthma);}$$

$$\frac{\sqrt{v(x_{(bm)})}}{x_{(bm)}} = \text{relative standard error for } x_{(bm)} \text{ (i.e., youth at risk);}$$

$$\frac{\text{COV}(y_{(bm)}, x_{(bm)})}{\bar{y}_{(bm)} \bar{x}_{(bm)} n} = \text{relative covariance between } y_{(bm)} \text{ and } x_{(bm)};$$

$\bar{y}_{(bm)}$ = benchmarked national average for injury or asthma per farm from the survey;

n = the number of farms from the NASS sampling frame used to derive the covariance between $y_{(bm)}$ and $x_{(bm)}$.

Because the relative covariance between the estimated number of injury or asthma cases and the number of youth at risk is typically negligible, the estimate $v(R)$ may be approximated as:

$$v(R) = 1000^2 R^2 \left[\left(\frac{\sqrt{v(y_{(bm)})}}{y_{(bm)}} \right)^2 + \left(\frac{\sqrt{v(x_{(bm)})}}{x_{(bm)}} \right)^2 \right] \quad (\text{D.6.2})$$

Equation D.6.2 was the method used to derive the standard errors for all rates in this document.