Agreement of 3 carcass rinse sampling methods (split carcass, repeat rinse, and adjacent pair) on the detection of *Salmonella* contamination in broiler carcasses

Karen D. Galarneau,∗ R. Hartford Bailey,† and Robert W. Wills‡

∗Quest Animal Health and Biosecurity, 15609 Peace Valley Rd., Brookwood, AL 35444; and †Department of Pathobiology and Population Medicine, College of Veterinary Medicine, Mississippi State University, 240 Wise Center Dr., P.O. Box 6100, Mississippi State 39762

ABSTRACT Whole carcass rinse is the most common method used to determine *Salmonella* prevalence in broiler carcasses. However, there is a need to determine the carcass rinse sampling method that best measures the *Salmonella* status of a broiler carcass as it proceeds through processing, thus allowing the assessment of efficacy of interventions to meet Food Safety Inspection Services (FSIS) performance standards.

In this study, 3 paired carcass rinse sampling methods, namely split-carcass method (rinises of 2 halves of one carcass), repeat rinse method (rinse and rerinse of same carcass), and adjacent pair method (rinises of 2 adjacent carcasses), were evaluated during actual operations in commercial poultry processing plants in the southeastern United States. The purpose of the work was to determine which method resulted in greatest agreement of *Salmonella* status on paired broiler carcass rinses.

The adjacent pair method showed moderate agreement consistently in 3 trials of 150 pairs per trial with kappa values of 0.46, 0.55, and 0.46. The repeat rinse method showed substantial kappa agreement (0.64) in one trial and moderate kappa agreement (0.47, 0.41) in 2 other trials. In one trial, the repeat rinse method showed a significant difference in prevalence rates between repeated rinses. Even though the split carcass method showed moderate kappa agreement (0.58, 0.45) in 150 carcasses in each of 2 trials, the disadvantages of the split carcass method were that it was more labor and time intensive and the product was damaged, when compared to the other 2 methods. Overall, although prevalence estimates were fairly consistent between pairs by each method, agreement between *Salmonella* status of the paired samples was less than desired, mostly moderate. This lack of agreement should be considered in the design of studies assessing the efficacy of interventions for the control of *Salmonella* in broilers to meet FSIS performance standards.

Key words: *Salmonella*, agreement, kappa, broiler carcass, carcass rinse

INTRODUCTION

Performance standards conducted at poultry processing plants for the detection of *Salmonella* were developed to reduce the health risk to consumers posed by *Salmonella* contamination of broiler carcasses. The efficacy of the performance standards relies on accurately measuring the *Salmonella* status of the broiler carcasses being sampled. The whole carcass rinse (WCR) procedure is the most common method used to determine *Salmonella* status in broiler carcasses and has been widely used in studies of *Salmonella* prevalence in broiler carcasses (Cason et al., 1997; Cox et al., 1983; Izat et al., 1991; James et al., 1992a, 1992b; Jetton et al., 1992; Jimenez et al., 2002; Lillard, 1989).

The statistical power of studies measuring microbial intervention efficacy in poultry systems could be increased if the same carcasses were measured before and after the treatment was applied rather than relying on changes in prevalence in groups of broilers. The use of paired carcass halves (one half for control and the other for treatment) was suggested by Izat et al. (1990) as a possible means of reducing variability in trials designed to test the effectiveness of various chemicals in reducing or eliminating *Salmonella* within processing waters. In their studies, which looked at most probable number (MPN) estimates of *Salmonella* in the carcass halves that were spiked with the pathogen, they found that the variation between the 2 halves of a split carcass was equal and significantly less than the variation among individual whole carcasses (Izat et al., 1990). Cason and Berrang (2002) tested the findings of Izat et al. (1990)
on *Salmonella* in carcass halves by comparing the MPN of bacteria on paired broiler carcass halves from broiler carcasses taken from a commercial processing plant at the point just before the chiller. They concluded that paired carcass halves had the advantage of reduced variability in numbers of aerobic bacteria, *Campylobacter*, *Escherichia coli*, and other coliforms over counts on different carcasses.

Carcass-to-carcass variability in *Salmonella* numbers in WCR studies has been reported by several authors (Cason and Berrang, 2002; McNab et al., 1993; Renwick et al., 1993). Renwick et al. (1993) investigated the variability and determinants of carcass bacterial load from WCR of roaster chickens. The samples were collected from the evisceration line of a commercial poultry processing plant in Canada over a 5 month period, and the lot and supplier were identified. Carcass-to-carcass variability accounted for 73.2% of the total variability in bacterial loads of prechill carcasses. The remainder of the variability in bacterial load was attributed to between-lots-within-supplier (14.2%) and between-supplier (12.6%) variability, respectively.

The comparison of repeat rinses from a broiler carcass has been studied with conflicting results (Izat et al., 1991; Lillard, 1989). In a study on broiler carcasses from the processing line of a commercial processing plant, Lillard (1989) found that *Salmonella* was not always recovered from the initial carcass rinse but was recovered from subsequent carcass rinses of the same broiler carcass. In a study of variability in broiler carcass bacterial load, a series of 5 repeat rinses were conducted on one bird from each of 96 study lots (McNab et al., 1993). Results showed a positive association between the bacterial count of the first rinse and the subsequent rinses. Regression analysis showed that carcasses with high counts on the initial rinse continued to have high counts when compared to carcasses with low initial counts. Also observed was a declining trend in the counts in the subsequent rinses (McNab et al., 1993). On the other hand, in a study done to determine the accuracy of a single WCR in estimating the MPN of *Salmonella*, no significant difference was found in 4 consecutive carcass rinses of the same broiler carcass that had been inoculated with *Salmonella typhimurium* (Izat et al., 1991).

Although most of the aforementioned studies used the MPN and quantitative methods, current Food Safety Inspection Services (FSIS) performance standards at poultry processing plants still rely on carcass rinses that detect the presence or absence rather than the enumeration of *Salmonella*. It should also be noted that there is a need for studies at commercial processing plants that are conducted during actual operations. There have been studies on the prevalence of *Salmonella* but not on agreement of *Salmonella* status of carcass rinse samples collected in a commercial plant during normal operations. The goal of this study was to measure agreement between the presence or absence of *Salmonella* in rinse samples obtained using 3 different sampling strategies (split carcass, repeat rinse, and adjacent pair).

**MATERIALS AND METHODS**

Five trials were conducted at commercial poultry processing plants in the southeastern United States. Two trials were conducted to test agreement for the split carcass rinse method, and 3 trials were conducted to determine agreement for both the adjacent and the rerinse carcass rinse methods.

**Split Carcass Rinse Method**

For the split carcass rinse method, 150 broiler carcasses were removed, approximately equally spaced over the course of 1.5 h in each trial, from the processing line after the inside-outside spray wash but before the final rinse cabinet. Using fresh latex gloves, carcasses were placed in individual sterile plastic bags, placed in a cooler, and immediately taken to a place at the plant off the processing floor. Each carcass was aseptically divided in the midline using a fresh pair of autoclaved shears, and each half (A and B) was placed in an individual sterile plastic bag with 100 mL Butterfield solution and shaken for one min according to established procedure (Cox et al., 1983; Izat et al., 1990). The split carcass rinse samples were collected in individual sterile Nalgene (NalgeNunc International, Rochester, NY) containers and transported on wet ice to the laboratory for *Salmonella* isolation.

**Repeat Rinse and Adjacent Pair**

To evaluate the repeat rinse (rinse and rerinse) and the adjacent pair (adjacent carcass rinse) sampling methods, 150 pairs of broiler carcasses were sampled using WCR in each of the 3 trials. Using fresh latex gloves for each carcass, two adjacent carcasses were removed from the processing line after the inside-outside spray wash but before the final rinse cabinet. Each broiler carcass was then placed in an individual sterile plastic bag with 100 mL Butterfield solution and shaken for one min as previously described (Cox et al., 1983). The first of each pair of carcasses was returned to the line after the rinse sample was collected. Its rinse sample was labeled with its pair number and an “adjacent” designation. Following its first rinse, the second of each pair of carcasses was placed in a second sterile plastic bag with Butterfield solution and subjected to a second WCR. The first and second rinses of this carcass were labeled by pair number and designated “rinse” and “rerinse,” respectively. After the second rinse sample was collected, the second broiler carcass was returned to the processing line. The rinse samples were collected in individual sterile Nalgene containers and transported on wet ice to the laboratory for *Salmonella* isolation.
**Salmonella Isolation**

For all trials, the combination of tetrathionate (TT) (Remel Inc., Lenexa, KS) broth and Rappaport-Vassiliadis (RV) broth selective enrichment method as described by Rybolt et al. (2004) was used to isolate *Salmonella*. Approx. 50 to 75 mL rinse was recovered from each broiler carcass rinse. The rinses were brought up to 1× buffered peptone water (BPW) with the addition of 10× BPW. The sample was mixed and then incubated for 24 h at 42°C. One mL rinse sample was transferred to a tube containing 9 mL TT broth. The tube was vortexed and incubated at 42°C for 48 h. From the tube, 0.1 mL was transferred to a tube containing 9.9 mL RV broth. The tube was vortexed and then incubated at 42°C for 18 to 24 h. A loopful of RV broth was streaked on an XLT4 agar plate and incubated at 37°C for 18 to 24 h. *Salmonella*-like colonies were confirmed biochemically using triple sugar iron (TSI) agar (Difco Laboratories, Detroit, MI) and lysine iron agar (LIA) (Difco Laboratories); they were confirmed serologically using anti-*Salmonella* poly A-I and Vi serum (Difco Laboratories).

**Sample Size**

For each of the carcass rinse method trials, 150 pairs of broiler carcasses or split-carcass halves were collected. Power analysis using Cantor’s sample size equation (Cantor, 1996) estimated that for each trial, the sample size of 150 pairs would allow the discrimination between kappa values for 2 methods of 0.4 and 0.6 assuming an α level of 0.05, power of 0.80, and prevalence estimates of 0.50 by either method.

**Statistical Analysis**

Assessing agreement and differences in prevalence for the different methods was performed using the FREQ procedure of SAS for Windows version 9.3 (SAS Institute, Inc., Cary, NC). The Tables/Agree statement generated the kappa coefficient of agreement between adjacent and rinse samples for the adjacent pair method, between rinse and rerinse samples for the repeat rinse method, and between half A and half B for the split carcass method within each of the appropriate trials.

The kappa coefficients were interpreted according to the categories suggested by Landis and Koch (1977): kappa less than 0 = poor agreement; 0.00 to 0.20 = slight; 0.21 to 0.40 = fair; 0.41 to 0.60 = moderate; 0.61 to 0.80 = substantial; and 0.81 to 1.00 = almost perfect. McNemar’s test was used to determine if the prevalence of *Salmonella* was different between adjacent and rinse samples for the adjacent pair method, between rinse and rerinse samples for the repeat rinse method, and between half A and half B for the split carcass method within each of the appropriate trials. A significance level of 0.05 was used for all analyses.

**RESULTS AND DISCUSSION**

**Split Carcass Method, Trials 1 and 2**

The prevalence and agreement results of each of the 5 trials are presented in Table 1. In trial 1, the prevalence of *Salmonella* was 81.3% in half A and 78.7% in half B; in trial 2, it was 63.3% in half A and 70.7% in half B. McNemar’s test was used to compare the prevalence between the paired samples, and no significant differences were found between the prevalence rates in half A and half B for trial 1 (P = 0.371) or Trial 2 (P = 0.071). This finding, however, does not necessarily mean that the *Salmonella* status of individual pairs of samples was the same.

In trial 1, the results showed that both halves of the split carcasses agreed (concordant pairs) that 20 (13.3%) were negative and 110 (73.3%) were positive for *Salmonella* contamination. There were 20 pairs (13.3%) of split carcass halves that did not agree (discordant pairs). In trial 2, both carcass halves A and B agreed (concordant pairs) that 31 (20.67%) were negative and 82 (54.67%) were positive. There were 37 discordant pairs (24.67%).

To measure the degree of *Salmonella* status agreement between the half A and half B samples, kappa coefficients were determined. There was moderate agreement according to categories from Landis and Koch (1977) between carcass halves in trial 1 (kappa = 0.58, P < 0.0001) and in trial 2 (kappa = 0.45, P < 0.0001).

---

**Table 1.** Results of kappa agreement, prevalence, and McNemar’s test for *Salmonella* status of split carcass, repeat rinse (rinse and rerinse), and adjacent pair (rinse and adjacent) rinse sampling methods.

<table>
<thead>
<tr>
<th>Trial</th>
<th>n</th>
<th>kapp*a</th>
<th>Prevalence</th>
<th>McNemar’s test</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Split carcass</td>
<td>Repeat rinse</td>
<td>Adjacent</td>
</tr>
<tr>
<td>1</td>
<td>150</td>
<td>0.58</td>
<td>81.3</td>
<td>78.7</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>150</td>
<td>0.45</td>
<td>63.3</td>
<td>70.7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>150</td>
<td>0.64</td>
<td>24.7</td>
<td>20.0</td>
<td>24.7</td>
</tr>
<tr>
<td>4</td>
<td>150</td>
<td>0.47</td>
<td>69.3</td>
<td>66.0</td>
<td>68.7</td>
</tr>
<tr>
<td>5</td>
<td>150</td>
<td>0.41</td>
<td>22.7</td>
<td>12.0</td>
<td>17.3</td>
</tr>
</tbody>
</table>

*aThe test of the null hypothesis that kappa = 0 was <0.0001 for all comparisons.
Repeat Rinses and Adjacent Broiler Carcasses, Trials 3, 4, and 5

The prevalence of *Salmonella* in trial 3 was 24.7% in the adjacent, 24.7% in the rinse, and 20.0% in the rerinse broiler carcasses. In trial 4, the prevalence was 68.7% in the adjacent, 69.3% in the rinse, and 66.0% in the rerinse broiler carcasses. In trial 5, the prevalence was 17.3% in the adjacent, 22.7% in the rinse, and 12% in the rerinse broiler carcasses.

There were no significant differences between the prevalence of *Salmonella* in the adjacent carcasses in trial 3 ($P = 1.000$), trial 4 ($P = 0.853$), or trial 5 ($P = 0.117$). Nor were there significant differences between the repeat rinses of the same broiler carcass in trial 3 ($P = 0.108$) or trial 4 ($P = 0.398$). However, there was a significant difference between the prevalence of *Salmonella* in the repeat rinses of the same broiler carcass in trial 5 ($P = 0.002$). This finding in trial 5 may be an effect of the first rinse removing some of the *Salmonella*.

In trial 3, the comparison between adjacent broiler carcasses showed that both broiler carcasses agreed (concordant pairs) that 22 (14.7%) were positive and 98 (65.3%) were negative for *Salmonella* contamination. However, there were 30 (20.0%) adjacent broiler carcass pairs that did not agree (discordant pairs). In trial 4, the comparison between adjacent broiler carcasses showed that both broiler carcasses agreed (concordant pairs) that 89 (59.3%) were positive and 32 (21.3%) were negative for *Salmonella* contamination. However, there were 29 (19.3%) adjacent broiler carcass pairs that did not agree (discordant pairs). In trial 5, the comparison between adjacent broiler carcasses showed that both broiler carcasses agreed (concordant pairs) that 17 (11.3%) were positive and 107 (71.3%) were negative for *Salmonella* contamination. However, there were 28 (17.3%) adjacent broiler carcass pairs that did not agree (discordant pairs).

Following categories from Landis and Koch (1977), there was moderate agreement for the adjacent broiler carcasses in trial 3 (kappa = 0.46, $P < 0.0001$), trial 4 (kappa = 0.55, $P < 0.0001$), and trial 5 (kappa = 0.46, $P < 0.0001$). One possible explanation for only moderate agreement is greater variation in *Salmonella* contamination status between adjacent broiler carcasses. However, this is contrary to the assumption that adjacent carcasses on the processing line have similar *Salmonella* contamination status.

It has been hypothesized that *Salmonella* cross-contamination of adjacent carcasses can occur during the inside-outside spray wash (Jimenez et al., 2002). However, the study done by Jimenez et al. (2002) in Argentina was based on an increase in *Salmonella* prevalence from carcasses visibly contaminated with feces at the shower stage compared to the prevalence from carcasses with visible fecal contamination at the evisceration stage, rather than a comparison of actual adjacent carcasses. On the other hand, in a study of broiler carcasses processed with an inside-outside bird washer, no cross-contamination of adjacent carcasses was found when uncontaminated broiler carcasses were placed adjacent to artificially contaminated carcasses during washing (Smith et al., 2005). The absence of cross-contamination was attributed to the use of a pilot size 30.5 cm shackle center separation distance as opposed to the standard used by industry of 15.3 cm. In this current study, the naturally occurring *Salmonella* status of broilers was determined in actual processing conditions, with standard industry shackle separation distance.

In trial 3, the comparison between the repeat WCR samples yielded 131 concordant pairs (24 pairs *Salmonella* positive and 107 pairs *Salmonella* negative) and 19 discordant pairs. In trial 4, the comparison between the repeat WCR samples yielded 115 concordant pairs (84 pairs *Salmonella* positive and 31 pairs *Salmonella* negative) and 35 discordant pairs. In trial 5, the comparison between the repeat WCR samples yielded 124 concordant pairs (13 pairs *Salmonella* positive and 111 pairs *Salmonella* negative) and 26 discordant pairs.

In trial 3, there was substantial agreement between repeat rinses (kappa = 0.64, $P < 0.0001$). This agreement was higher compared to that found in the adjacent pairs rinse method. There was moderate agreement between repeat rinses in trial 4 (kappa = 0.47, $P < 0.0001$) and trial 5 (kappa = 0.41, $P < 0.0001$). This agreement is lower compared to that found in the adjacent pairs rinse method for trials 4 and 5. The variation in agreement in repeat rinses compared to adjacent carcasses may be attributed to variation between multiple rinses of the same carcass. This type of variation between repeat rinses was also reported by Lillard (1989) in a study of commercially processed poultry carcasses at pre- and post-evisceration steps.

**Overall Comparisons**

Although prevalence can affect agreement as measured by kappa, wherein paradoxically, extremely high and extremely low prevalence would result in lower kappa coefficient, this was not observed in this study. Even though there was substantial variation in prevalence among trials, there were no extremely high or low prevalence estimates. Considering the results of the analyses, the highest and lowest kappa agreement was observed among repeat rinses of the same carcass. The highest, or substantial, agreement (kappa = 0.64) was observed in the repeat rinse method in trial 3; the least, or moderate, agreement (kappa = 0.41) was found also in the repeat rinse method in trial 5. On the other hand, the adjacent pairs rinse method showed a consistently moderate kappa agreement without any significant differences in prevalence.

The repeat rinses (rinse and rerinse) of the same carcass showed inconsistent agreement among 3 trials.
(ranging from moderate to substantial) and showed a significant difference in the prevalence between rinse and rerinse in trial 5. The significant difference in prevalence indicates that rinsing the same carcass twice may be washing off the Salmonella during the first rinse, resulting in the second carcass rinse being negative. These results show that the concern regarding the impact of repeat rinses on the Salmonella contamination status of the sampling unit is a valid consideration.

The split carcass method also showed moderate agreement. However, in practice, this method was found to present logistical issues requiring more time, sterile instruments, and labor than using the standard carcass rinse process.

The results of this study are based on data collected under commercial processing plant conditions during actual operations and using naturally contaminated samples. The broiler carcasses were picked at the area before the inside-outside bird wash and before chilling. However, it is equally important to note that the prevalences detected were not reflective of the final prevalences for the flocks processed in the plants because the rinses were done prior to chilling.

The use of the presence-absence method for this study was selected over enumeration to be more comparable to compliance testing for performance standards on Salmonella in poultry processing plants, which is currently based on the presence or absence of Salmonella in carcass rinse samples. Furthermore, the FSIS uses Salmonella isolation methods that measure the presence or absence of the organism on the carcass by enriching the carcass rinse and the selective culturing methods. This current work used a delayed secondary enrichment culture method (Rybolt et al., 2004) that increases the chance of isolating any Salmonella that may be present. Although enumeration methods potentially provide more information in regard to the level of contamination, they also generally suffer from a lack of sensitivity compared to enrichment techniques applied to the original rinse sample. At low levels of contamination, enumeration methods may lead to the carcass rinse being judged free of Salmonella, whereas enrichment methods, as used in this study, would allow detection of contamination even if the number of bacteria is not determined. Enumeration data allow a better assessment of the degree of change in magnitude of Salmonella numbers during processing; however, as long as performance standards rely on prevalence, the investigation of the effect of interventions on the presence or absence of Salmonella in carcass rinse samples remains relevant.

It is important to note that similar prevalence estimates between 2 populations do not necessarily mean true agreement. Overall, prevalence estimates were fairly consistent between pairs by each method, although there was evidence that the rerinse samples resulted in lower prevalence estimates compared to rinse or adjacent carcass samples. The moderate agreement between the Salmonella status of the paired samples for all 3 methods was less than desired. This relative lack of agreement should be considered in the design of studies assessing the efficacy of interventions for the control of Salmonella in broilers. The results point to the need for more sensitive and consistent sampling methods in determining the Salmonella status of broiler carcasses.

ACKNOWLEDGMENTS

This research was supported by USDA–National Integrated Food Safety Initiatives (USDA-NIFSI) award number 2006-51110-03645. Karen Galarneau is grateful to the Fulbright Philippines Fellowship Program and the International Institute of Education.

The authors are grateful to Terry Doler, Food Safety Laboratory, College of Veterinary Medicine, Mississippi State University, for leading the laboratory testing. They are also grateful to Mary Ann Ballard, Adam Pochocki, Kelly Hataway, Kevin Howe, and Sanaz Salehi for their efforts in sample collection.

REFERENCES


