

21 **Summary of results:**

22

- 23 1) Some establishments were able to produce, consistently, AMR product that did not
24 contain central nervous tissue (spinal cord or dorsal root ganglia), based on not
25 detecting central nervous tissue in 6 or more samples. On the other hand, some
26 establishments had nearly all tested samples positive for central nervous tissue,
27 suggesting that much of the AMR product from these establishments would contain
28 central nervous tissue.
- 29 2) For the study, approximately 35% of the finished AMR samples had central nervous
30 tissue detected; 29% of the samples had spinal cord tissue detected; and 10% had
31 dorsal root ganglia tissue detected. However, the percentages of positive samples
32 were significantly different for different periods of the survey. For the last third of
33 the survey, the percentage of positive (for any central nervous tissue) post-desinewing
34 samples was about 44%, versus 31% for the first two-thirds of the survey.
- 35 3) The occurrences of spinal cord tissue and dorsal root ganglia tissue are not
36 significantly correlated, suggesting that there may be different factors that cause their
37 occurrences.
- 38 4) Type of bones used in processing may contribute to the likelihood of central nervous
39 tissue being present in the finished product, however, regardless of the type of bone
40 used, product can be produced with a low likelihood of central nervous tissue being
41 present.
- 42 5) The average of the calcium values for post-desinewing samples is 91.7 mg/100g; the
43 highest value is 159 mg/100g. The average of the iron to protein ratios is 0.35; the
44 average of the excess iron (iron - 0.13(1.1) protein) measurements is 2.94 mg/100g.
- 45 6) Excess iron and calcium are positively correlated, suggesting a common set of factors
46 that influence their levels. In general, higher levels of these variables are associated
47 with a higher likelihood of central nervous tissue being in the product, suggesting that
48 these variables reflect processing parameters that might be related to the likelihood of
49 central nervous tissue being present in the product. However, there were significant
50 expectations to general trends.
- 51 7) Relationships between excess iron and calcium with machine operating parameters
52 and product type are equivocal.
- 53 8) Protein levels decreased, on average, by about 3%, as a result of the desinewing
54 step. Since the protein that is being removed is most likely iron-deficient, this finding
55 suggests that excess iron measure be adjusted to account for this loss of protein.
- 56 9) Thirteen percent of the post-desinewing samples that were negative for central
57 nervous tissue had positive finding for the matched pre-desinewing sample. This
58 suggests that the quantity of sample used for determining the presence of central
59 nervous tissue may be too small.
- 60 10) An Elisa procedure was compared with the direct procedure of detecting central
61 nervous tissue. While there is a correlation between the Elisa results and the findings
62 of central nervous tissue on samples by the direct method, there were also significant
63 number false negative findings. Using a cutoff value for determining positive samples
64 that provides an approximate 25% false positive rate, the false negative rate on the all
65 positive samples, as determined by the direct method, was about 30%; for samples
66 from establishments for which most of their samples were positive, the false negative

67 rate was about 20%. Of further concern though, false negative rates seem to be a
68 function of the establishment from which the samples are taken; no discernable
69 “reason” was found to help explain possible causes for this establishment –specific
70 dependency.

71 **11)** Twenty-three establishments that had positive spinal cord samples submitted 67
72 follow-up AMR samples for verification (which were not included in the survey
73 results). Most of the follow-up samples were collected after the survey. All, but two,
74 establishments had no more than 3 follow-up samples; the two exceptional
75 establishments had 12 and 19 samples. For these two establishments, the percentages
76 of positive survey and follow-up samples were nearly the same, of about 40-50%.
77 For the other 21 establishments the percentages of positive follow-up samples,
78 collected after the survey was completed, were generally less than the percentages for
79 the survey samples. Overall, approximately 1/3 of the follow-up samples were
80 positive for spinal cord tissue.

81 **Analysis of results of survey of AMR bovine product derived from vertebra.**

82
83 The purpose of this survey is to characterize AMR product derived from beef vertebra, in
84 particular, to establish a baseline for the prevalence of central nervous tissues (cnt),
85 consisting either of spinal cord (sc) and dorsal root ganglia (drg), in AMR products.
86 These two types of tissues, have been identified in the Harvard BSE Risk Assessment
87 (2002) as specific risk materials for bovine spongiform encephalopathy (BSE). The
88 results of this survey are to be used for a regulatory impact analysis, as required by
89 Congress.

90 A primary question is what processing factors affect the likelihood of central
91 nervous spinal tissue and dorsal root ganglia being found in samples of AMR product. In
92 addition, FSIS is concerned about the presence of excessive iron (exFe) in the product, as
93 evidence of the presence of more than negligible amounts of bone marrow tissue, and
94 calcium, as a measure of minutely sized bone particles that are not normally seen in hand-
95 deboned beef. In a proposed regulation (1998), FSIS proposed requirements regarding
96 the level of iron in AMR product that is considered to be excess over what would exist in
97 the corresponding hand-deboned product, as a means of evaluating whether more than
98 negligible bone marrow exists in the AMR product. In Attachment A is an analysis of
99 iron levels in the hand deboned data from the 1996 FSIS AMR survey. Based on this
100 data, it was derived that an excess iron measure of 2.8 mg/100g or more for a single
101 sample would imply that the sample was, with more than 99.9% confidence, not from a
102 hand -deboned product processed under good manufacturing practices. This conclusion
103 was based on an analysis of variance of the excess iron levels for the hand-deboned
104 product and took into account the analytical measurement error associated with iron and
105 protein. Thus, the identification of the factors that influence the levels of excessive iron
106 and calcium are also desired. Variables, for which information was collected, are:
107 machine pressure and dwell times during processing; types of bones (rib, pelvis, flat, or
108 any other types); type of vertebra (neck or back only); and food chemistry (iron, protein
109 and calcium).

110 Data on product before and after desinewing were collected. The regulatory
111 (present and proposed) requirements apply to the AMR product after desinewing;
112 however, FSIS is interested in the relationships of the product before and after
113 desinewing.

114 Thirty-four establishments were identified as producing AMR beef product. The
115 sample design called for 6 pairs of samples to be collected, one sample before and one
116 after desinewing for each AMR machine operating in the establishment. The times of
117 sampling were from the middle of January, 2002 to the end of August, 2002, with one
118 sample being received for analysis on September 18, 2002. However, no samples were
119 collected during the last week of February and the month of March. If an establishment
120 was not producing product at the time that the sample was requested, an additional
121 sample request form was sent to that establishment. The follow-up sample request form
122 apparently was the cause of there being more than 6 samples in some establishments;
123 samples for the original form and also for the additional form were collected.
124 Furthermore, cost and resource constraints prevented the planned designed to be
125 completed. Toward the end of the survey period, the designated pre-desinewing samples

126 were not collected. In addition, for some of the later samples, food chemistry analyses
127 were not done. Hence the distribution of samples over the establishments is not balanced.

128 Pathological examination for the presence of spinal and dorsal root ganglia
129 nervous tissue is time consuming, expensive and requires an expert technician. These
130 burdens motivated the development of an Elisa procedure that would be quick and simple
131 to conduct.

132 Since the data is collected from establishments over a 6 month period, it is not
133 possible to identify causal relationships of variables, as would or might be the case in a
134 controlled study. Identified correlations between two variables can occur as a result of
135 unknown causal variables affecting both. Furthermore, variables that are causal, in the
136 sense that changes in values of a variable x affect the values of y, given everything else
137 being equal or held constant, may not be detected because of the existence of other,
138 unknown, variables that affect values of y, thus masking the causal relationship. Among
139 establishments, there are many variables that could affect the values of variables of
140 interest, thus masking relationships that exist, or creating correlations of variables that are
141 not truly related in a causal fashion. In an attempt to eliminate some of the effects of
142 unknown establishment-specific variables, statistical analyses of relationships are
143 performed within establishments, and summaries of these results are used to determine
144 the existence of possible causal relationships. But even within establishments, there exist
145 variables that could affect the results and possibly mask true relationships or create non-
146 causal correlations. In addition, in an attempt to eliminate the effect of possible deviant
147 or extreme results, non-parametric statistical tests are relied upon for evaluating the
148 strength of relationships. Thus the statistical analysis consists of examinations of the
149 consistency of relationships within establishments, and among establishments. Statistical
150 significant levels that are quoted are two-sided. Statistical analyses were performed using
151 PC-SAS[®], release 8.0.

152 An attachment presents an analysis of data from the FSIS, 1996 AMR survey of
153 beef derived from neckbone, where samples were collected from 7 establishments
154 producing AMR and two establishments producing hand-deboned product. The purpose
155 of the analysis was to derive an excess iron performance standard for AMR product.

156

157 **General overview of results**

158

159 Before presenting a more detailed analysis of the relationships of the variables that were
160 studied in this survey, summary results, by establishment, are presented in Table 1.
161 Included in Table 2 are: the fraction of samples that were positive for either spinal cord
162 (sc) or dorsal root ganglia (drg), designated by "cnt," as well as the fraction that are
163 positive for each one; the mean of the pressures and dwell times used for the collected
164 samples; and the median calcium and excess iron result. The exact formula for the excess
165 iron will be given below. The determination of the fraction of samples that are positive
166 with respect to the specific type nervous tissue or any of them is as follows: if any of the
167 matched pre-or post- desinewing samples were positive for a nervous tissue, then the
168 sample (considered as the pair) was counted as a positive. The fraction then is just the
169 ratio of the number of positive samples divided by the number of samples, where in the
170 case the sample consists of a matched pair it is counted as one sample. The median values
171 for the food chemistry variables are the results on the post-desinewing sample. The

172 establishment identification numbers (id) were determined by assigning random numbers
 173 from 1 to 34 to the establishments. The observations in Table 1 are ordered from the
 174 smallest to the largest fractions of cnt positive samples, and, within groups that had the
 175 same fraction, by the median excess iron result.
 176

177 Table 1: Summary of results for selected variables, by establishment. Observations are
 178 ordered in ascending fraction of samples detected with central nervous tissue (spinal cord
 179 or dorsal root ganglia).
 180

estab- lish- ment id	mach- ine type	numb. samples	mean pressure (psi)	mean dwell time (s)	frac. cnt +	frac. sc +	frac. drg +	numb. food chem. samples	median calcium (mg/100g)	median excess iron (mg/100g)
28	P	7	1658	1.1	0.00	0.00	0.00	5	95.0	1.69
33	H	6	2902	4.2	0.00	0.00	0.00	5	54.0	2.14
16	H	13	3300	25.5	0.00	0.00	0.00	8	74.0	2.67
32	H	6	2757	3.0	0.00	0.00	0.00	4	109.5	3.05
17	H	11	2233	4.2	0.09	0.00	0.09	5	58.0	1.14
8	H	7	2443	1.0	0.14	0.00	0.14	5	63.0	1.35
15	P	7	1223	1.0	0.14	0.14	0.00	4	116.0	2.36
3	H	7	2621	3.0	0.14	0.14	0.00	5	97.0	2.76
26	H	12	2853	4.0	0.17	0.00	0.17	8	84.0	2.99
30	H	6	2921	7.1	0.17	0.00	0.17	2	88.5	4.04
6	H	5	3135	3.5	0.20	0.20	0.00	2	116.0	3.69
34	H	12	3429	1.0	0.25	0.17	0.08	7	88.0	3.58
27	P	6	2322	1.8	0.33	0.00	0.33	5	105.0	3.61
5	H	6	2902	3.9	0.33	0.17	0.17	5	93.0	4.39
20	P	6	1451	0.1	0.33	0.33	0.00	4	50.0	1.05
11	H	6	2500	5.0	0.33	0.33	0.00	5	66.0	1.54
18	H	6	2757	2.8	0.33	0.33	0.00	4	102.0	3.82
23	P	10	2345	2.3	0.40	0.20	0.20	5	121.0	2.71
1	P	7	2467	4.0	0.43	0.29	0.29	5	132.0	3.85
22	H	7	3213	1.0	0.43	0.43	0.14	4	90.5	4.80
9	H	11	3004	4.6	0.45	0.45	0.00	8	74.0	1.61
14	P	6	2878	1.0	0.50	0.17	0.33	5	91.0	2.67
25	H	6	3000	1.0	0.50	0.50	0.00	5	100.0	2.75
7	H	6	2983	10.0	0.50	0.50	0.00	4	97.0	4.27
21	P	6	2909	0.5	0.67	0.33	0.50	4	87.0	2.55
19	H	6	2612	5.0	0.67	0.67	0.00	5	89.0	3.11
13	H	6	2868	2.8	0.67	0.67	0.00	4	99.5	3.16
12	H	6	2394	2.0	0.67	0.67	0.33	5	118.0	3.22
24	H	6	2612	1.6	0.67	0.67	0.00	4	123.5	4.97
31	H	5	2840	1.0	0.80	0.80	0.00	2	104.0	3.42
2	H	16	3376	3.9	0.81	0.50	0.31	12	96.0	3.06
10	H	6	2902	1.0	0.83	0.83	0.17	4	119.0	4.06
4	H	13	3104	1.0	0.85	0.85	0.08	8	94.0	2.02
29	H	7	2600	10.0	0.86	0.86	0.43	3	65.0	1.93

221 An examination of the results presented in Table 1 reveals that: 1) there is
 222 virtually no correlation between the fraction of spinal cord (sc) positive and dorsal root
 223 ganglia (drg) samples; 2) there are moderate correlations - determined by significance
 224 levels ranging from about 0.10 to 0.25 for Spearman correlations of approximately 0.2 in
 225 absolute value - between the fraction of positive central nervous tissue (cnt) with the
 226 median excess iron, calcium, and the mean machine pressure and dwell times; 3) there is
 227 virtually no correlation of the fraction of positive drg samples and the variables identified
 228 in 2), and 4) there are high positive correlations (P- value < 0.05) of the median excess
 229 iron with median calcium and machine pressure. The lack of correlation of the fraction
 230 of sc and drg tissues could imply that distinct factors contribute to the likelihoods of sc
 231 and drg tissue.

232 Samples were classified regarding the likelihood that they would be positive with
 233 respect to cnt tissue by considering the fraction of samples detected to have cnt tissue for
 234 the establishment from which they were sampled. As depicted in Table 1, five

235 classifications were made: 1) establishments for which less than 10% of the samples were
236 positive for cnt; 2) establishments for which the fraction cnt positive samples is greater
237 than 0.10, but less than or equal 0.25; 3) establishments for which the fraction cnt
238 positive samples is greater than 0.25, but less than or equal 0.50; 4) establishments for
239 which the fraction cnt positive samples is greater than 0.50 but less than 0.8; and 5)
240 establishments for which the fraction of cnt samples is greater than or equal to 0.8. For
241 class 1, there are 5 establishments; for class 2, 7 establishments; for class 3, 12
242 establishments; for class 4, 5 establishments; and for class 5, 5 establishments. To
243 distinguish positive cnt samples from negative ones, for a sample, a value of $\frac{1}{2}$ was added
244 to the class level if the sample were positive for cnt tissue.

245
246

247 **Analysis of AMR product characteristics after desinewing.**

248

249 *Nervous tissue*

250

251 Since the regulation is to apply to product after desinewing, and the number of samples
252 after desinewing is nearly balanced with respect to machines, the following analysis is on
253 the post-desinewing samples. There were 256 results from samples collected after
254 desinewing. Of these there were 28.9% of them were detected having sc tissue, 9.8%
255 having drg tissue, and 96.5% having hemeopathatic cells. The percentage of samples
256 positive for any central nervous tissue (cnt) is 35.2%. These numbers are consistent with
257 that premise that occurrences of drg and sc tissues are not correlated, since, if it is
258 assumed that the occurrences of sc and drg tissues are independent, there would be an
259 expected 35.9% of the samples that would be positive for either tissue. Within-
260 establishment correlations of the occurrences of spinal cord and dorsal root ganglia were
261 computed for the post-desinewing product. Of the 9 within-establishment correlations
262 that were possible to compute, 6 were positive and 3 were negative, which is not
263 statistically significant (P- value of about 0.5).

264

265 A further analysis of the data revealed that over the time of the survey the
266 percentage positive samples increased. For the first two months of the survey (Jan-Feb.
267 2002) 31 samples were collected. The sampling was interrupted during the month of
268 March, and continued from April to the end of August. The samples were divided into 6
269 time periods: the first period consisting of samples received by the laboratory during
270 January and February; and the other 5 periods were assigned by dividing the samples into
271 periods of approximately equal numbers, assuring that the samples arriving at the lab on
272 the same dates were in the same period (the actual collection dates were not recorded for
273 all samples). The following table presents the fraction of positive post-desinewing
274 samples for any central nervous tissue, by time period and likelihood of positive samples,
275 defined from Table 1. The result presented in Table 2 clearly show that the percentages of
276 positive samples generally increase over time for all likelihood classes. The latter third
277 of the survey generally consisted of samples collected during the summer, thus, it is
278 possible that environment, specifically temperature, may have an impact on the
279 likelihood of samples being positive for central nervous tissue.

279

280

281 Table 2: Fractions of post-desinewing samples that were positive, by likelihood classes
 282 of positive samples and time periods of survey. The lower time period designation means
 283 the samples were analyzed earlier in the survey.

284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310

```

..ffffffff..ffffffffffffffffffffffffffffffffffffffffffffffffffffffff..fffffffffff+
likelihood class
,   1   2   3   4   5   All
,   1   2   3   4   5   All
,   frac. , frac. , frac. , frac. , frac. , frac. ,
,   N   cnt   N   cnt   N   cnt   N   cnt   N   cnt   N   cnt
,   period
,   1   6, 0.000, 5, 0.000, 11, 0.273, 4, 0.500, 5, 0.600, 31, 0.258,
,   2   7, 0.000, 9, 0.222, 20, 0.300, 6, 0.500, 6, 0.500, 48, 0.292,
,   3   7, 0.000, 9, 0.222, 14, 0.143, 10, 0.700, 8, 0.750, 48, 0.354,
,   4   7, 0.000, 11, 0.091, 14, 0.429, 2, 0.500, 8, 0.625, 42, 0.310,
,   5   10, 0.000, 10, 0.300, 10, 0.400, 5, 1.000, 11, 0.818, 46, 0.457,
,   6   6, 0.167, 11, 0.091, 14, 0.429, 3, 0.667, 7, 1.000, 41, 0.415,
,   All 43, 0.023, 55, 0.164, 83, 0.325, 30, 0.667, 45, 0.733, 256, 0.352,
$ffffffff<fff<ffffffff<fff<ffffffff<fff<ffffffff<fff<ffffffff<fff<ffffffff<fff<fffff&
    
```

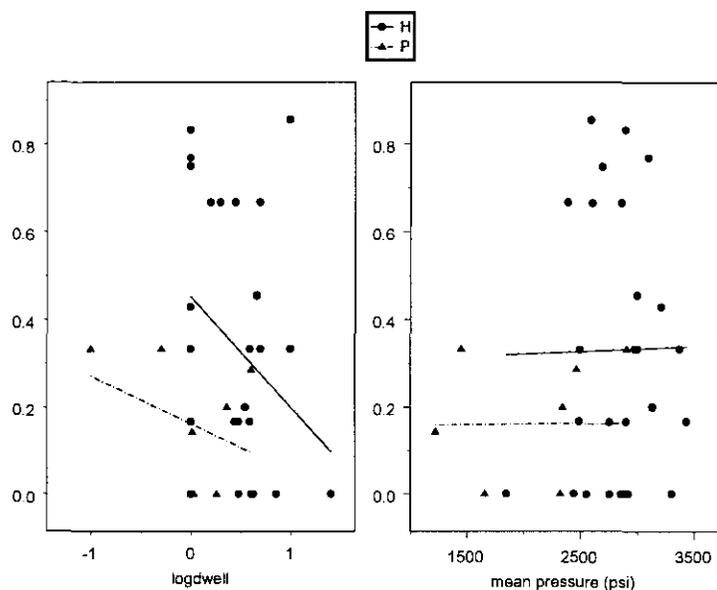
311 A further analysis (not presented here) indicated that the pattern seen above in Table 2
 312 did not hold for the presence of dorsal root ganglia tissue, and (consequently) did hold for
 313 the presence of spinal cord tissue. The existence of a time factor effect would not have a
 314 significant impact on the conclusions, regarding information collected for establishments,
 315 insofar as the samples, within establishments, over time are nearly balance; the variation
 316 associated with this variable is not large compared to that associated with the
 317 establishment facto; and the machine settings (pressure and dwell time) are not
 318 significantly correlated with time, though for the dwell time there were some
 319 establishments for which the dwell times were higher in the latter third of the survey.
 320 Because food chemistry results were not collected for the latter third of the survey and
 321 the percentage of positive samples were higher in the latter third of the survey, the
 322 estimated relationship of the food chemistry results and the likelihood of positive samples
 323 derived from these data may not be accurate. The time effect is included in the
 324 subsequent analyses by defining a variable to be equal 0 when the sample is within the
 325 first two-thirds of the survey, and equal 1 when the sample is in the last third of the
 326 survey.

327
 328 Two types of AMR machine, Protocon (P) and Hydrosep (H), were identified in
 329 the 34 establishments producing the AMR product: 26 were using Hydrosep machines
 330 and 8 were using Protocon machines. The means of the establishment-specific percentage
 331 of positive cnt samples were: for the Hydrosep, 37.7%; for the Protocon, 30.5%. An
 332 analysis of variance on the variable with value equal 1 when cnt was detected, and 0
 333 otherwise, assuming establishment as a random factor, and including a time effect
 334 distinguishing samples within the first two-thirds of the survey from the latter third, did
 335 not indicate significant machine type effects. In addition, machine effects, within
 336 establishments, were not statistically significant for the H machines, but marginally so for
 337 the P machines (P- value = 0.11).

338 The difference of the percentages of cnt positive samples for the two machine
339 types is not statistically significant. However, the pattern of the differences of the
340 percentages of sc and drg positive samples for the two machines are dissimilar. For the
341 Hydrosep machines, the establishment-specific mean of sc positive samples is 34.5%
342 compared to 16.2% for the Protocon. This difference is significant at about the 0.19 level,
343 based on analysis of variance on the raw results, treating establishment as a random
344 factor, and accounting for the time effect. The Wilcoxon test statistic (on the
345 establishment-specific mean values) had a significance level of 0.16. For the presence of
346 the drg tissue, the direction of the machine effect is reversed: For the Hydrosep machines,
347 the drg establishment specific percentage is 7.6% compared to 16.1% for the Protocon
348 machines (P- value = 0.13 for the ANOVA, and 0.30 for the Wilcoxon). The difference
349 of the differences is significant at the 0.05 level, based on an ANOVA, with
350 establishment as a random factor and the time effect. This finding reinforces the premise
351 that different factors are contributing to the presence of the two types of central nervous
352 tissue. For the presence of homeopathic cells, there was no significant difference, where
353 the percentage of positive samples is about 97% for the H machine, and 94% for the P
354 machine.

355 An important question is whether there is a relation between the percentage of sc
356 or drg tissue positive samples and values of operating parameters for the machine:
357 pressure and dwell time. One establishment had two machines that had pressures that
358 were more than slightly different, so in the following analysis, these machines were
359 considered as two establishments. The averages of these for a given machine within an
360 establishment were assigned to the machine and to samples for which the information
361 was missing. The averages of the machine pressures and dwell times for the Hydrosep
362 machines are slightly higher than those for the Protocon machines. The higher pressures
363 and dwell times might provide a “possible reason” for the higher fraction of positive CNS
364 tissues and the lower fraction of positive drg tissues for the Hydrosep machine samples.
365 Figure 1 presents a scatterplot of the establishment-specific fraction of samples with
366 spinal cord tissue versus the \log_{10} of the mean dwell time and the mean pressure, with
367 different symbols indicating the type of machine. Also the linear regression lines for the
368 two machines are depicted.

369
370 Figure 1: Scatterplot of fraction of establishment specific positive spinal cord tissue
371 samples versus \log_{10} mean dwell time (s) (left graph) and mean pressure, (right graph),
372 together with linear regression lines, by type of machine. The y-axes are the fraction of
373 samples that are positive for sc tissue; the x-axes are labeled.
374



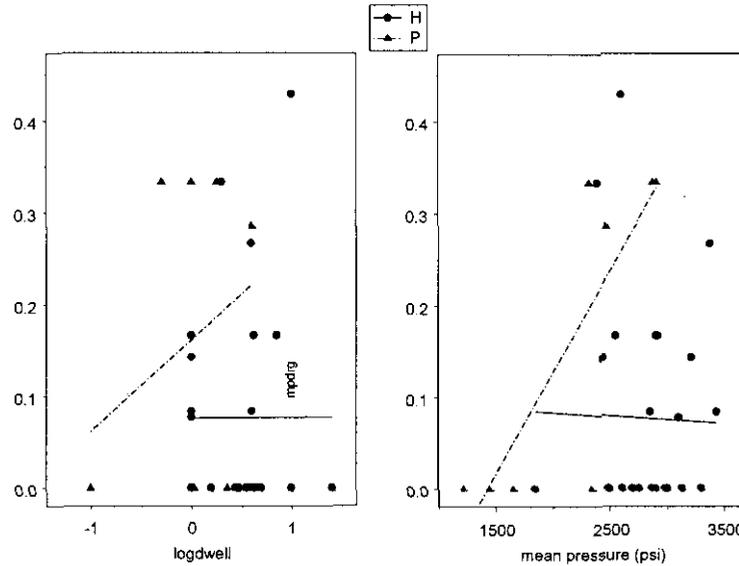
375
376
377

378 For the mean pressure, the lines are nearly parallel to the x-axis indicating the lack of
379 correlation of the likelihood of sc tissue and machine pressure. For the dwell time, the
380 correlation appears to be negative: higher likelihoods of sc tissues are associated with
381 lower dwell times.

382 A similar examination of the effects of pressure and dwell time on the presence of
383 dorsal root ganglia yielded no consistent results. Figure 2 presents scatterplots and linear
384 regression of lines of the establishment-specific fraction of samples detected positive for
385 drg tissue versus the \log_{10} of the dwell time and the machine pressure. While a positive
386 correlation is seen for the Protocon machines, none is seen for the Hydrosep machines.
387

388 Figure 2: Scatterplot of fraction of establishment specific positive drg tissue samples
389 versus \log_{10} mean dwell time (s) (left graph) and mean pressure, (right graph), together
390 with linear regression lines, by type of machine. The y-axes are the fraction of samples
391 that are positive for drg tissue; the x-axes are labeled.
392

393



394 To examine further the relationship of these variables with the presence of central
 395 nervous tissue, relationships within the establishments are examined. To help decide the
 396 significance of any pattern, the following approach is used. Assume random variables, x
 397 and y , measured on a sample, where x takes on values of 0 (for a negative sample) and 1
 398 (for a positive sample) and where y is a variable that can take on any numerical value. To
 399 determine if there is a correlation between x and y , the average of the ranks of the values
 400 of y for positive samples is compared to the average rank, $(n+1)/2$, where n is the
 401 number of samples. Specifically, the statistic computed for each establishment is
 402

403
$$d_k = \left(\bar{r}_{m_k} - \frac{n_k + 1}{2} \right) \delta_k \quad (1)$$

404
 405 where the index k specifies an establishment, m_k is the number of positive samples out of
 406 n_k samples of the establishment, $\delta_k = m_k/n_k$ is the fraction of positive samples, \bar{r}_{m_k} is the
 407 average rank of the of the m_k positive samples among the n_k values of y (the lowest value
 408 being assigned the lowest rank of 1, and the next lowest a rank value of 2, and so forth,
 409 and ties are set equal to the average rank). This statistic is symmetric about $\delta = 1/2$. Note
 410 that d_k is zero when $\delta_k = 0$ or 1, or when all the rank scores of y are the same. The
 411 variance of d_k , when the null hypothesis of zero correlation is true, assuming no ties, is
 412
 413

414
$$\text{var}(d_k) = (n_k + 1) \delta_k (1 - \delta_k) / 12. \quad (2)$$

415
 416 The test statistic computed is
 417

418
$$T = \sum_{k=1}^K n_k d_k \quad (3)$$

419

420 where K is the number of establishments. The variance of T is

421

422
$$\text{var}(T) = \sum_{k=1}^K n_k^2 \text{var}(d_k). \quad (4)$$

423

424 Observations for which there were no differences in the rank values were deleted. Hence,
425 to gauge the significance of the value of T for testing whether there is a relationship
426 (rejecting the null hypothesis of no relationship), a z-value is computed,

427

428
$$Z = \frac{T}{\sqrt{\text{var}(T)}} \quad (5)$$

429

430 which is compared to the percentiles of the normal distribtuion.

431

432 None of the 6 values of Z, of Eq. 5, for determining the significance of
433 comparisons: the presence of central nervous tissue (cnt), spinal cord (sc) and dorsal root
434 ganglia (drg), with dwell times and machine pressure, were statistically significant at the
435 two-sided 0.10 level or less. The six values are: $Z(\text{cnt, dwell}) = 0.667$; $Z(\text{sc, dwell}) =$
436 1.30 ; $Z(\text{drg, dwell}) = -0.873$; $Z(\text{cnt, press}) = 0.287$; $Z(\text{sc, press}) = 1.53$; $Z(\text{drg, press}) = -$
437 1.61 . The number of establishments for which there were non-zero d_k 's was small: for
438 the drg and dwell time, only 4 values were non-zero, and all were negative. The sign for
439 the correlation between spinal cord and dwell time was positive, which is not the same as
440 the sign for the between-establishment correlation of these seen above (Fig. 1). Of some
441 interest though is the difference of the signs of the Z-values for the sc and drg, again
442 reaffirming the possibility of different factors affecting the presence of spinal cord and
443 dorsal root ganglia nervous tissue.

443

444 *Product type used in processing*

445

446 One of the factors that might affect the likelihood of sc or drg tissue is the type of
447 product used in processing, in regards to the vertebra or other bones that are in the pre-
448 processed product. Information was collected on whether rib or pelvis or other types of
449 bones were used; another field recorded whether neck or back vertebra were used, as
450 opposed to just the whole vertebra (which might have included neck or back vertebra).
451 For each establishment, the fraction of samples from product that were processed using
452 neck or back vertebra only, and included bones was computed. For the most part, within
453 establishments, added bones were not used (less than 20% of the samples) or most of the
454 time added bones were used (greater than 60%).

455

456 For designated neck or back vertebra, one clear pattern was found: all high
457 likelihood of cnt establishments (classes 4 or 5) had very few samples that indicated only
458 neck or back vertebra were used. For added bones, a similar pattern can be seen, where
459 only 2 of the 10 high likelihood establishments indicated no added bone usage. However,
there were a few establishments with high fractions of added bones and non-specified

460 vertebra that had a low fraction of samples with cnt (likelihood class 1 or 2), thus, no
 461 general causal relationship involving these variables can be established from these data.
 462 The number of samples for vertebra designated as only neck of back is small, thus it is
 463 difficult to make conclusions regarding the effect of only using neck vertebra bones.
 464 Further the samples not designated as using only neck or back vertebra could, in any
 465 case, consist mostly of neck or back vertebra, adding to the difficulty of inferring possible
 466 causal relationships involving this variable. Table 3 gives the fraction of cnt positive
 467 samples by machine type, vertebra type and whether other bones were added or not.

468
 469 Table 3: Summary of fraction of samples with cn tissues, added bone and vertebra types,
 470 machine type, and level of dwell time (defined in table and in text).

likelihood cnt class	machine type	estab- ishment id	fraction cnt	fraction non- specified vertebra	fraction added bones	mean pressure (psi)	mean dwell time (s)
471 1	H	17	0.00	0.20	0.00	1850	4.24
472 1	H	17	0.17	0.33	0.17	2552	4.23
473 1	H	32	0.00	1.00	0.00	2757	3.00
474 1	H	16	0.00	1.00	0.08	3300	25.54
475 1	H	33	0.00	1.00	1.00	2902	4.17
476 1	P	28	0.00	0.57	0.00	1658	1.07
477 2	H	3	0.17	0.33	0.17	2491	3.00
478 2	H	8	0.14	0.71	0.86	2443	1.00
479 2	H	26	0.08	0.92	0.92	2853	4.00
480 2	H	6	0.20	0.40	1.00	3135	3.52
481 2	H	34	0.25	1.00	1.00	3429	1.00
482 2	H	30	0.17	1.00	1.00	2921	7.14
483 2	P	15	0.14	0.00	0.00	1223	1.03
484 3	H	25	0.33	1.00	0.00	3000	1.00
485 3	H	5	0.17	0.50	0.17	2902	3.88
486 3	H	18	0.17	1.00	1.00	2757	2.75
487 3	H	7	0.33	1.00	1.00	2983	10.00
488 3	H	11	0.33	0.83	0.00	2500	5.00
489 3	H	22	0.43	1.00	1.00	3213	1.00
490 3	H	9	0.45	1.00	0.91	3004	4.65
491 3	P	14	0.33	0.67	1.00	2878	1.00
492 3	P	23	0.20	0.00	0.00	2345	2.29
493 3	P	27	0.33	0.83	0.67	2322	1.80
494 3	P	1	0.43	0.86	0.00	2467	4.00
495 3	P	20	0.33	0.17	0.83	1451	0.10
496 4	H	12	0.67	1.00	0.00	2394	2.00
497 4	H	19	0.67	1.00	0.67	2612	5.00
498 4	H	13	0.67	1.00	0.67	2868	2.83
499 4	H	24	0.67	1.00	0.83	2612	1.58
500 4	P	21	0.67	0.50	0.33	2909	0.51
501 5	H	29	0.86	0.86	0.00	2600	10.00
502 5	H	31	0.75	0.75	0.75	2700	1.00
503 5	H	10	0.83	0.83	0.83	2902	1.00
504 5	H	4	0.77	0.92	0.92	3104	1.00
505 5	H	2	0.60	0.93	0.93	3374	3.91

515 *Food Chemistry*

516
 517 Protein (p %), iron (Fe mg/100g) and calcium (Ca mg/100g) measurements were made on
 518 170 of the 256 samples. For each sample, a single determination was made, which may
 519 in itself, increase the variability of the iron results since the repeatability of iron
 520 measurements is about 0.16 mg/100g (see attachment A). Thus, a 95% confidence
 521 interval for the true level in the sample based on a single analysis would have a range of
 522 0.64 mg/100g. However, this estimate was based on an analysis of data for which a
 523 handful of "outlier" results were deleted. In practice, it is recommended that duplicate

524 analysis be made, and the mean value used as an estimate of the concentration of iron in
525 the product, provided the individual results are not “too” far apart.

526 For each sample, the iron to protein ratio (ipr) was computed and an excess iron
527 measurement was determined as: $ex\ Fe = Fe - (0.13)(1.1)p$. The factor 0.13 is the ipr
528 value determined from ground product collected in the FSIS AMR 1996 survey,
529 representing product that is used for AMR processing. The factor 1.1 is an adjustment
530 (or fudge) factor to account for the loss of iron-depleted protein during desinewing (See
531 the attachment and the next section). The actual factors are not critical for this analysis;
532 what is important is the relationship of iron with the other variables.

533

534 Iron versus protein

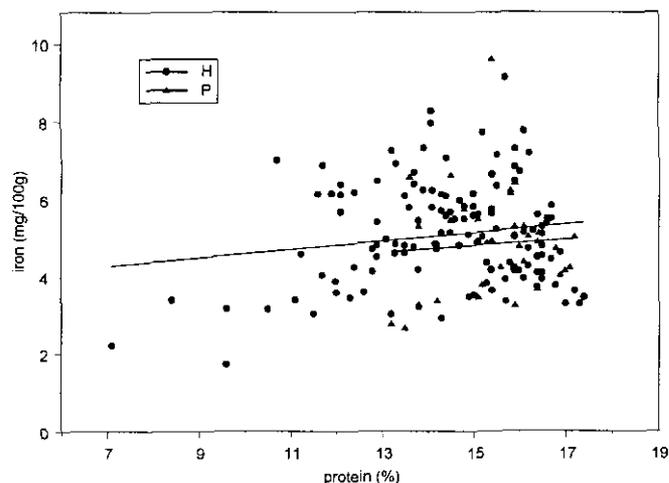
535

536 The assumption that iron levels are correlated with protein levels in hand-deboned
537 beef provides the motivation for using excess iron or the iron-to-protein ratio as a
538 measure of product quality with regards to the evaluation of whether there are more than
539 negligible amounts of bone marrow in the product. This assumption was valid for the
540 hand-deboned data of the 1996 survey referred to above. If excess iron were present in
541 the AMR product, then the correlation for AMR product would be smaller, in so far as
542 added bone marrow that is the assumed cause of the added iron would not contribute to
543 the protein levels. Thus, for a high(er) protein product, high(er) levels of iron would be
544 expected, but the percentage increase due to the addition of bone marrow would be less
545 than that of a low(er) protein product. Hence, for AMR product, the lack of a significant
546 correlation does not in itself represent evidence that invalidates the assumption. For these
547 data, there is not a large correlation when computed over all samples. Figure 3 is a
548 scatterplot of iron versus protein, by machine type, together with the linear regression
549 lines.

550

551 Figure 3: Scatterplot of iron versus protein, by machine type, with linear regression lines.

552 The shorter line is the linear regression line for the data from the Protocon machines.



553 For the 170 samples, the Pearson correlation of iron and protein is 0.13, and of excess
554 iron and protein, equal to -0.09 . However, the non-parametric correlations (Spearman
555 and Kendall) of iron and protein were nearly zero, and those of excess iron and protein,
556 were negative.

557 A fairer evaluation of the hypothesis of positive correlation between iron and
558 protein, given everything else being equal, is based on the correlations that exist within
559 each establishment and machine. For each machine, the Spearman correlations were
560 computed and were for the most part positive: 30 of the 41 correlations computed were
561 positive, 1 was zero, and 10 were negative. This pattern is significant at approximately
562 the 0.0139 level, based on the sign test. For the Pearson correlation, the results are
563 similar: for the sign test the significance is 0.05. When the same statistical test is
564 computed using the excess iron measurement instead of the iron measurement, the
565 significance level is 0.76 for the sign test. For the Spearman correlation the significance
566 level is 0.14. The conclusion is that iron is correlated with protein.

567

568 Excess iron and calcium

569

570 The mean of the calcium results for the Hydrosep machine is about 89 mg/100g;
571 for the Protocon, it is about 100 mg/100g. For iron, there was one reported result at 15.7
572 mg/100g; the next highest result was 9.6 mg/100g. The 15.7 mg/100g result is
573 considered as an outlier and thus deleted from the analyses. The mean of the excess iron
574 results for the Hydrosep machine is about 2.89 mg/100g; for the Protocon, it is about
575 2.47 mg/100g. As will be discussed below, these differences are not statistically
576 significant. First, the relationship of excess iron and calcium and of their relationships
577 with the likelihood of cnt in samples is explored.

578 Pearson and Spearman correlations of the machine-specific median calcium and
579 excess iron values were computed over all machines and found to be about 0.6 for both,
580 indicating that calcium and excess iron results are positively correlated across machines.
581 Figure 4 is a scatterplot of excess iron versus calcium, together with linear regression
582 lines for data from the two machine types. It is seen from the graph that the excess iron
583 levels on the average are higher for the data from the Hydrosep machines; the calcium
584 levels are nearly the same, and the positive correlations of excess iron and calcium for
585 data from the two machines are about the same. This latter result suggests that there are
586 production factors that may have significant deleterious influence on both calcium and
587 excess iron results.

588 The highest value for calcium is 159 mg/100g, with a corresponding excess iron
589 value of 4.1 mg/100g. The sample came from a class 4 establishment, a Hydrosep
590 machine, and the sample was positive for spinal cord tissue. The second highest calcium
591 result is 150 mg/100g, from a class 3 establishment, Hydrosep machine, however, the
592 excess iron result for that sample is 2.9 mg/100g, which is about the average, and the
593 sample was negative for cnt.

594 The highest excess iron result (excluding the one sample with iron result of 15.7
595 mg/100g) is 7.23 mg/100g and the calcium value is also a relatively high at 123 mg/100g.
596 This sample came from a class 2 establishment, a Protocon machine, with a positive
597 central nervous tissue sample. For these establishments, most of the samples were
598 negative for cnt, however, this one particular sample was not. The second highest excess

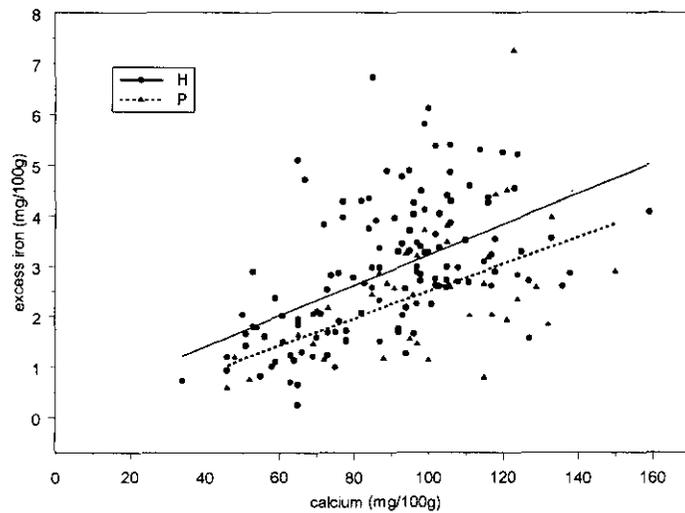
599 iron result is 6.7 mg/100g, from a class 3 establishment, Hydrosep machine, however, the
600 calcium result for that sample is a moderate 85 mg/100g and the sample was negative for
601 cnt. These examples suggests that calcium and excess iron, taken together, could provide
602 a clue regarding the likelihood of positive cnt samples, and thus, the quality of
603 processing.

604 This latter notion is buttressed somewhat by examining the relationship of excess
605 iron and calcium by the likelihood of cnt tissue classification. In Table 4 is presented the
606 median excess iron and calcium levels over samples within likelihood classes, by
607 machine. The excess iron and calcium entries in the table are highly correlated. Of
608 particular interest are the results for the class 1 samples for the Hydrosep machine: the
609 median excess iron result is 2.13 mg/100g and the median calcium result is 68 mg/100g.
610 These are both the lowest of those presented associated with the Hydrosep machines.

611

612 Figure 4: Scatterplot of excess iron versus calcium, with linear regression lines for data
613 from the Hydrosep and Protocon of machines.

614



681	2	8.40	3.36	0.40	70.00	2.08
682	2	13.70	6.67	0.49	111.00	4.59
683						

684

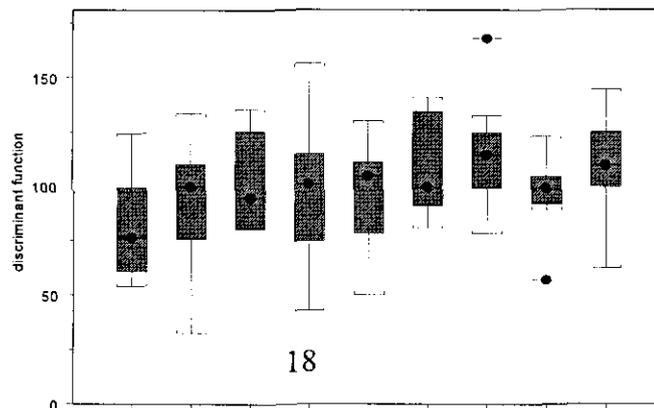
685 The results in Table 5 are interesting but also perplexing. All the above samples, except
 686 one (the 4th row), were produced from product for which the type of vertebra used was
 687 not specified, and all samples were from product that used bones of an unspecified type
 688 (not ribs, pelvis, or flat bones but others). The high iron results, excess iron results, and
 689 calcium results, all produced on machine 2, are associated with high protein results. In
 690 the following analyses, the two machines for this establishment of Table 5 were
 691 considered as separate establishments.

692 The 6 values of Z (Eq. 5), corresponding to the relations of excess iron and
 693 calcium with the occurrences of any central nervous tissue (cn), spinal cord (sc), and
 694 dorsal root ganglia (drg) tissues are: $Z(\text{exfe, cn}) = 0.30279$; $Z(\text{exfe, sc}) = -0.22916$;
 695 $Z(\text{exfe, drg}) = 0.89923$; $Z(\text{ca, cn}) = 1.91766$; $Z(\text{ca, sc}) = 0.85937$; and $Z(\text{ca, drg}) =$
 696 1.79846 . These results show that, within establishments, calcium is statistically
 697 significantly related with the likelihood of central nervous tissue being present in the
 698 samples. The values of excess iron are less so related, which could in part be due to the
 699 higher relative error of the estimated values due to sampling and analytical procedures.

700 A discriminant analysis was performed to help develop a simple function that
 701 could be used to distinguish between the positive and negative cnt samples. From the
 702 analysis, a discrimination function, $f = 6\ln(\text{exFe}) + \text{ca}$, was determined. As is evident,
 703 the value of calcium has more influence than that of excess iron, since typically the
 704 values of $\ln(\text{exfe})$ are between -0.5 and $2 \ln(\text{mg}/100\text{g})$, whereas the range of calcium is
 705 34 to $159 \text{ mg}/100\text{g}$. Figure 5 gives boxplots of the values of the discriminant function, by
 706 likelihood of cnt classes, defined above. The distribution of the function's values is
 707 "lower" for the class 1 samples, reinforcing the observation made above concerning the
 708 lower levels of iron and calcium for samples from establishments that had all negative
 709 samples. If the criterion were that the value of the function is greater than 100 , then the
 710 false positive rate for classification of samples from class 1 would be about 30% . The
 711 false negative rate for the high likelihood classes (4 and 5) for this rule would be
 712 approximately 22% ; 29% for class 3 and 60% (3 of 5 samples) for class 2. Clearly a
 713 criterion based on individual sample values of calcium and iron alone could not be used
 714 to discriminant between samples with and without cnt tissue, however, the analysis shows
 715 that a profile of these values may reflect general quality of processing, which in turn
 716 affects the likelihood of positive cnt product.

717

718 Figure 5: Discriminant function: $6\ln(\text{exfe}) + \text{ca}$, for classifying samples containing central



719 nervous tissue (cnt), by likelihood classes of cnt, defined in text.

720

721 An analysis of variance is performed on the values of the discriminate function. First the
 722 results of the analysis for the $\ln(\text{exFe})$ and calcium are given. For calcium, there is a
 723 relatively small between machine, within-establishment variance component compared to
 724 that of excess iron: for calcium, the intra-machine correlation is about 10%, whereas for
 725 excess iron, it is about 40% and for the natural log of excess iron, it is about 58%. The
 726 within establishment standard deviation (computed as the square root of the within and
 727 between machine variance components) is about 14 mg/100g for calcium, 1.1 mg/100g
 728 for the excess iron and about 0.5 $\log(\text{mg}/100\text{g})$ for the natural log of the excess iron).
 729 Relatively speaking, the variation of excess iron is larger than that of calcium. For the
 730 values of the discriminant function, $f = 6\ln(\text{exFe}) + \text{ca}$, the intra-machine correlation is
 731 estimated to be about 28.6%, and the total within establishment variance is 266.7; thus
 732 the standard deviation of f (for a single analysis of a sample) is 16.3.

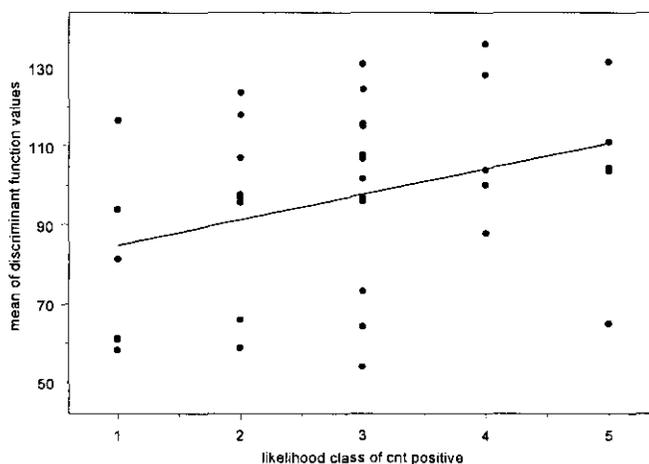
733 However, as can be seen from Figure 5, there are many values of the discriminant
 734 function that are relatively low. Figure 6 presents scatterplot of the establishment-
 735 specific mean values of the discriminant function versus the likelihood of positive cnt
 736 class, together with a regression line, treating the likelihood class variable as a
 737 independent variable (just for visual demonstration purposes with no intention of
 738 prediction). While the figure depicts the “trend” of the mean values, the exceptions are
 739 also clearly shown.

740

741 Figure 6: Establishment –specific mean values of discrimination function versus
 742 likelihood of positive cnt samples classes. Regression line shows trend, but is not meant
 743 for prediction.

744

745



746

747

748 Relations of excess iron and calcium with processing parameters

749

variable										
ca	0.115	1.000	0.072	0.424	1.000	-0.456	0.154	1.000	0.005	
dwell	0.178	0.030	-0.124	0.160	-0.462	0.128	0.241	-0.070	-0.101	
pbone	0.417	0.111	-0.001	-0.192	-0.192	-0.495	0.266	0.086	-0.135	
pressps1	-0.078	0.296	-0.037	-0.170	0.010	0.108	-0.132	0.217	-0.018	
vet	0.049	-0.129	-0.103	-0.366	-0.473	0.063	-0.204	-0.352	-0.020	

Note that calcium is positively correlated with excess iron, but excess iron is negatively correlated with machine pressure while calcium is positively correlated with machine pressure (as indicated in the analysis of covariance discussed above). Of the 21 Spearman correlations between excess iron and dwell times, 4 are negative (1 of which belonged to a P machine) and 16 are positive, for a two-sided significance (ignoring machine type) of 0.01 for the sign test. Of the 27 Spearman correlations of excess iron and machine pressure, 18 are negative and 8 are positive, for significance of 0.076 for the sign test. Of the 27 Spearman correlations of calcium and machine pressure, 6 are negative (4 belonging to the H machines), and 20 are positive, for significance of < 0.01 for the sign test. Of the 21 Spearman correlations between calcium and dwell times, 12 are negative and 8 are positive, for a two-sided significance (ignoring machine type) of 0.5 for the sign test. Of the 14 Spearman correlations of calcium and vertebra type, 8 are negative and 3 are positive, for a significance of 0.23 for the sign test. This analysis suggests that, within establishments, excess iron is positively correlated with dwell time and negatively correlated with machine pressure, and calcium is positively correlated with machine pressure.

Across establishments, the same pattern and significance of these correlations are not seen. The median establishment specific excess iron is negatively correlated (Spearman) with dwell times for the H machines (P- value = 0.22), and positively correlated for the P machines (P- value = 0.08); the median establishment specific calcium has nearly zero correlation with mean machine pressure for both P and H machines, a slight positive correlation with vertebra type for the H machines, but negative with the P machines. For the H machines, the correlation of the median establishment-specific calcium and dwell times was negative (P-value <0.01), but it was positive for the P machines. And the median excess iron is positively correlated with the means of the establishment specific machine pressures for both types of machines.

One further analysis was performed. Mixed linear effect models with the dependent variable of calcium and excess iron, assuming establishments and machines within establishments are random factors, were performed where log₁₀ dwell times for excess iron and machine pressure for calcium were considered as covariates, as above. For calcium, when adding covariates: vet and the interaction of vet and pbone, the test statistic equal to minus 2 times the loglikelihood ratio statistic decreased by about 20, which, based on 3 degrees of freedom, is statistically significant. Table 7 presents the mean calcium and excess iron levels for the types of products and machines by type of product.

Table 7: Mean calcium and excess iron levels for types of machines and product.

..fffffffff--fffffffffffffffffffffffffffffffff--fffffffffffffffff+

Machine Type											
H P All											
ca				exfe				ca			
N				Mean				N			
0	0	13	80.23	2.30	14	111.79	2.66	27	96.59	2.48	
1	3	95.00	2.54	5	66.60	1.30	8	77.25	1.76		
1	0	31	91.13	2.66	9	102.78	2.56	40	93.75	2.64	
1	1	86	90.22	3.09	9	94.44	2.79	95	90.62	3.06	

872 As is evident, there does not appear to be a consistent pattern; the statistical significance
 873 of the interaction seen in the mixed linear effect model does not translate to a practical
 874 significance. Thus, no conclusion or statement is being made concerning the possible
 875 effects of these factors on calcium or excess iron levels.

876 In conclusion, there seems to be evidence that would suggest, at least as a
 877 hypothesis, that, given everything else being equal, calcium is an increasing function of
 878 machine pressure and excess iron is an increasing function of dwell time. This is based
 879 on comparison within establishments. However, across establishments, these
 880 relationships do not hold. The uncontrolled nature of this study and the high variability
 881 of the results preclude developing estimates of functions that can be used to predict
 882 relationships; further, more controlled studies, are needed.

885 *Comparison of pre- and post-desinewing.*

887 There are 135 pairs of matched samples collected from pre- and post- desinewing. The
 888 desinewing operation remove cartilage and bone material from the product, which would
 889 include iron-depleted protein and calcium, thus it would be expected that calcium and
 890 protein levels would decrease, the iron to protein ratio and excess iron measure would
 891 increase. In addition, the desinewing operation would cause cns and drg tissue to occur
 892 more often.

893 In expectation, the above relationships were valid. However, measurement error
 894 and other factors created a significant number of comparisons that were in the opposite
 895 direction than expected. Of particular importance in the relationship of protein, pre- and
 896 post - desinewing, since the protein that is removed is iron-depleted, thus, causing, if all
 897 things remained equal, an increase in the iron to protein ratio from that that would be
 898 seen if the desinewing operation was not performed. For protein, 30% of the results had
 899 the pre-desinewing protein result lower than the post-desinewing result. The analytical
 900 standard deviation for protein is a function of the true level, $p, sd(p) = 0.03p^{0.65}$ (Price, et
 901 al, 1994), so that, for example, the standard deviation of measured values on a sample
 902 with a true protein level of 16% would be 0.17%. The difference of 2 independent results
 903 thus would be 0.25%. Hence, a positive difference greater than 0.50%, would, assuming
 904 normality of errors, with 97.5% confidence, represent paired samples with the true pre-
 905 desinewing protein value lower than that of the post-desinewing value. Ten percent of
 906 the samples had values had pre-desinewing protein measured values more than 0.5% less

907 than that of the paired post-desinewing values. These “unexpected” results could be due
908 to sampling error: the samples themselves might not have been good “representative”
909 samples of the product.

910 The means of the protein values pre- and post-desinewing for all the paired
911 samples are 15.1% and 14.7%, respectively, for an average difference of 0.42%. The
912 Pearson and Spearman correlations of the differences versus the pre-desinewing protein
913 values were significant at significance level of 0.075 and 0.20, respectively, whereas, the
914 correlations for the logarithmic transformed values were significant at the 0.81 and 0.37
915 levels. Thus to characterize the differences, the ratio of the post- to the pre-desinewing is
916 considered

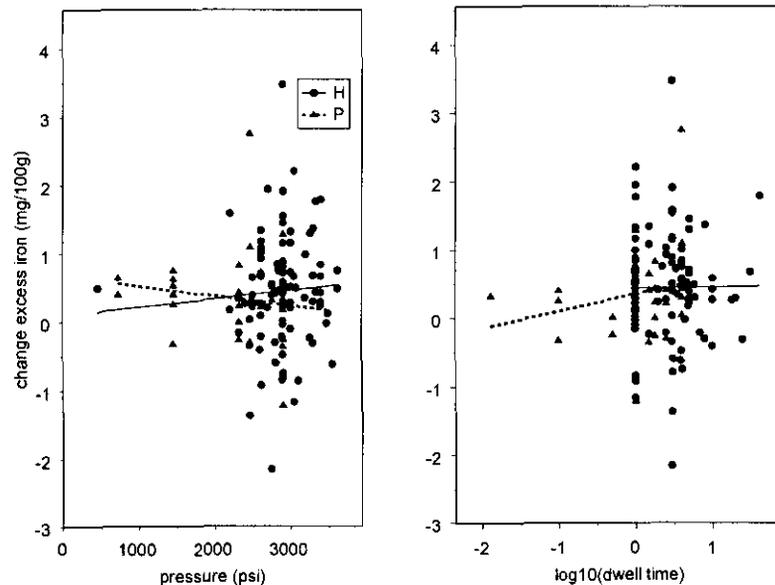
917 There was not a statistically significant machine effect, when assuming
918 establishment is a random factor, though there is a significant establishment effect (P-
919 value < 0.01). For the H machines, the geometric mean of the ratios of the post- to pre-
920 desinewing protein values is 96.5%, while that for the P machines is 98.5%. Over all
921 results, the geometric mean of the ratios is 96.9%. Thus, on average, through the
922 desinewing step, the protein content was reduced by 3%.

923 Figure 7 contains scatterplots of the change of excess iron versus the machine
924 pressure and \log_{10} of the dwell time. The plots indicate little correlation of the change of
925 excess iron with the machine pressure and dwell time.

926

927 Figure 7: Scatterplot of the change of excess iron versus the machine pressure and
928 \log_{10} of the dwell time, and linear regression lines, by type of machine.

929



930

931

932

933

934

The correlation between the change in levels of excess iron and the excess levels
in pre-desinewing product within establishments is negative. On average, the excess iron

935 level for the post-desinewing product was larger by 0.43 mg/100g than that of the pre-
936 desinewing product. The iron levels increased by about the same amount: from an
937 average of 4.76 mg/100g to an average of 5.20 mg/100g. Within-establishment
938 Spearman correlations of the changes in excess iron and the excess iron levels in the pre-
939 desinewing product were almost all negative (P-value < 0.001 for the sign test). However,
940 the median establishment specific changes in excess iron were not negatively correlated
941 with the median establishment - specific excess iron for the pre-desinewing product.

942 ANOVA models for change of excess iron were performed, with pre-desinewing
943 excess iron as the covariate, machine type as a fixed factor and establishment as a random
944 factor and included as an interaction effect with the covariate. The machine type effect
945 was significant at about the 0.05 significance level, and the “average” slope of the
946 covariate of pre-desinewing excess iron levels was negative and significant at better than
947 the 0.01 level. Based on these models, the average standard error of predicted increase in
948 excess iron from the pre- to post-desinewed product is about 0.44, comparable to the
949 mean increase of 0.43 stated above. Covariates of dwell times and machine pressure were
950 added to the model, but there was no significant improvement. Thus, while there does
951 appear to be a relationship of the change in the excess iron levels with that of the excess
952 iron levels in the pre-desinewed product, the relative error of the prediction of the amount
953 of the decrease is large.

954

955 Comparison of central nervous tissue pre- and post- desinewing.

956

957 There are 134 matched samples for which analyses of central nervous tissue were made.
958 Of these 134 samples, 94 of them were negative for central nervous tissue. However, of
959 these 94, 13%, or 12 of them were positive for the matched pre-desinewing sample; the
960 rate applies to both types of machines: 3 of 14 samples from the Protocon machines were
961 positive and 9 of 70 samples from the Hydrosep machines were positive. While the
962 number of samples is too small to discern any pattern, one establishment had 3 of them.
963 For this establishment 81% of the samples tested were positive. For the spinal cord
964 tissue, 7.4% of the 94 samples were positive, and for the dorsal root ganglia 5.7% of the
965 samples were positive.

966

967

968

969 **Analysis of ELISA procedure for determining the presence of cns tissue.**

970

971 A short explanation of the ELISA procedure is presented before the analysis of the
972 results. The Elisa is a sandwich immunoassay that utilizes two antibodies to GFAP [one
973 antibody is bound to the bottom of the wells, which entraps the GFAP from the sample, a
974 second antibody conjugated to peroxidase is then used to detect the bound GFAP].
975 Samples for analyses were formed by first taking enough material from the 1 pound of
976 product that was sent to the laboratory to form four blocks (1/2 x 3/4 x 1 inch each) of
977 tissue. These were placed in a small plastic bag and mixed as thoroughly as possible by
978 manipulating the bag. This comminuted product was then sampled by inserting a cotton
979 swab 3 times into the product at different sites. Excess material is removed from the
980 swab. Supposedly the swab will entrap, on average, about 50 mg of a meat sample. The

981 swab is then inserted into 1 ml of sample diluent, agitated to dislodge the meat sample
982 (20 times), and 50 ul of the 1000 ul of diluent/meat sample is added to one well. The
983 controls are freeze-dried bovine brain at four standard dilutions of "risk material"--
984 GFAP-- in buffer.

985

986 There were 295 Elisa measurements. The reported results were transformed by the
987 natural logarithm, because doing so provides a better resolution of the data, and because
988 the standard deviations of the logarithmic transforms of the results would be more
989 homogeneous than those of the untransformed results, since the repeatability standard
990 deviation of optical density (OD) responses is often an increasing function of the
991 expected OD. The natural log of the Elisa results reported as zero were assigned a value
992 of -12 , since the lowest non-zero Elisa transformed result was -11.5 . For the remainder
993 of the report, Elisa results will sometimes refer to the logarithmic transformed results;
994 from the context it should be clear which is meant. Analysis of variances and graphical
995 examination revealed no, or very little, significant differences between the non-zero Elisa
996 transformed responses between the pre- and post desinewed samples. Thus, this
997 designation will be ignored, at least initially, in the subsequent analyses.

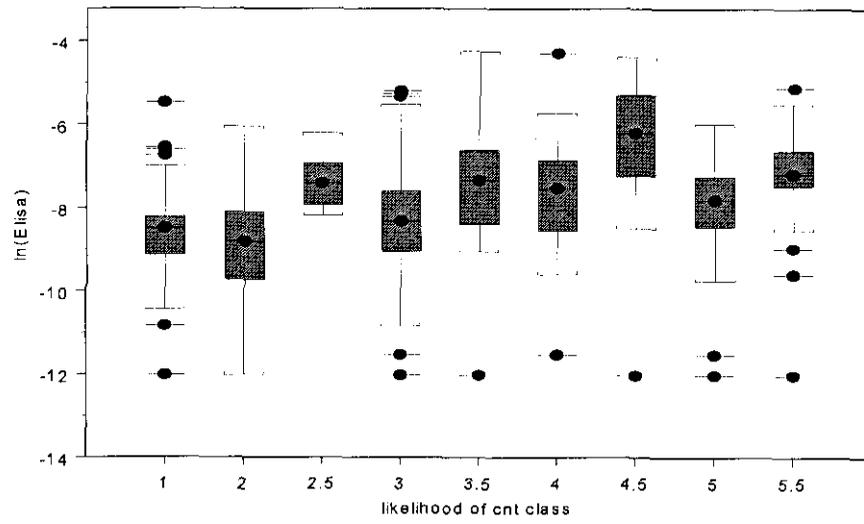
998 Discrepancies between results of the Elisa tests and the direct cn tissue
999 determination could be caused in part by sampling and measurement errors: matched or
1000 paired samples may actually be different with respect to the presence of cn tissue, or there
1001 may be amounts of cn tissue that are below the (direct) method's sensitivity contributing
1002 to a false negative result. Even if the tests for cn tissue were negative, the corresponding
1003 matched sample might contain sufficient amounts of cn tissue to cause an OD Elisa
1004 response, or vice versa. Thus, the comparison of the Elisa results with the detection of
1005 cnt in samples should account for the likelihood that the sample would contain cnt, even
1006 if it were not found by the direct method. As described above, 5 classes of samples were
1007 defined, depending upon the establishment's percentage of samples for which cn tissue
1008 was detected. The class 1 samples are most likely to represent samples that are negative.
1009 Figure 8 provide boxplots of the Elisa results by likelihood classes of cnt tissue.

1010

1011

1012 Figure 8: Boxplots for natural log of Elisa results, by classes defining the
1013 likelihood of cnt being present in sample. Class 1 is the lowest likelihood, and class 5.5
1014 is the highest likelihood. The classes that are whole integers consists of matched samples
1015 that were negative for cns tissue.

1016



1017

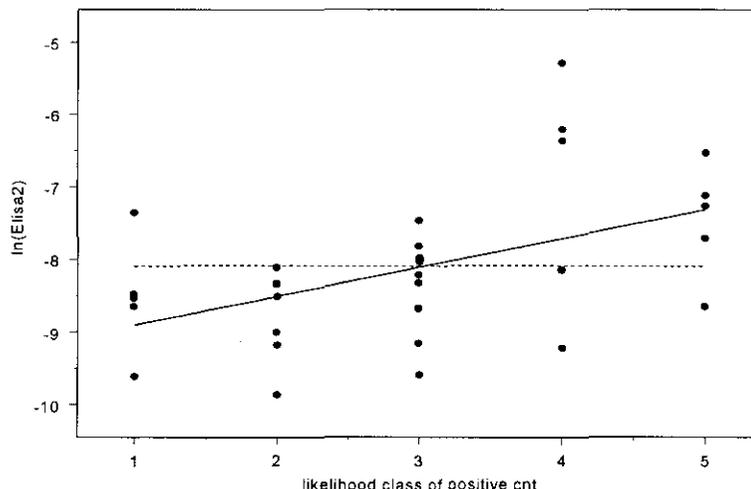
1081 Table 9: Results from 2 establishment, for which Elisa responses were different. All
 1082 samples represent post-desinewed product.

1083
 1084
 1085
 1086
 1087
 1088
 1089
 1090
 1091
 1092
 1093
 1094
 1095
 1096
 1097
 1098
 1099
 1100
 1101
 1102
 1103
 1104
 1105
 1106
 1107
 1108
 1109
 1110
 1111
 1112
 1113
 1114
 1115
 1116
 1117
 1118
 1119
 1120
 1121
 1122
 1123
 1124
 1125
 1126
 1127
 1128
 1129
 1130
 1131
 1132
 1133
 1134

Establishment id	calcium (mg/100g)	excess iron (mg/100g)	machine dwell time(s)	Positive for cnt direct method	Elisa result	Positive for cnt (>0.000304)
4	105.00	2.69	1.0	yes	.00078	yes
4	92.00	1.84	1.0	no	.	.
4	92.00	1.77	1.0	no	.00058	yes
4	82.00	2.19	1.0	yes	.00073	yes
4	94.00	2.28	.	no	.00074	yes
4	94.00	1.36	1.0	yes	.00075	yes
4	111.00	2.80	1.0	yes	.00123	yes
4	127.00	1.66	1.0	yes	.00421	yes
4	.	.	.	yes	.00167	yes
4	.	.	1.0	yes	.00615	yes
4	.	.	1.0	yes	.00260	yes
4	.	.	1.0	yes	.00414	yes
4	.	.	1.0	yes	.00093	yes
2	97.00	3.59	1.0	yes	.	.
2	95.00	3.44	2.5	no	.	.
2	105.00	2.70	1.5	yes	.	.
2	97.00	3.02	4.4	no	.	.
2	85.00	3.09	4.6	no	.00106	yes
2	95.00	3.43	5.0	yes	.00006	no
2	117.00	2.72	4.2	no	.00014	no
2	98.00	2.97	2.3	no	.00023	no
2	110.00	3.65	4.3	yes	.00024	no
2	94.00	2.67	4.4	yes	.00071	yes
2	95.00	5.00	5.5	yes	.00205	yes
2	83.00	2.79	4.3	no	.00022	no
2	.	.	5.2	yes	.00007	no
2	.	.	5.2	yes	.00000	no
2	.	.	4.3	yes	.00007	no

The results from Table 9 clearly depict possible problems with characterizing the performance of the Elisa procedure; the Elisa did well for establishment 4, in which all Elisa tested samples were positive. However, the Elisa did not do so well for establishment 2, in which only 3 of the 11 samples tested were evaluated as positive and 5 samples had false negative Elisa results relative to positive results for the direct method. Figure 9 gives the establishment-specific mean values of ln(Elisa) versus the likelihood of positive cnt class, together with a regression line (just to depict the relationship) and the derived cutoff value (dotted line) for distinguishing a positive result, based on the 25% rule developed above.

Figure 9: Establishment-specific mean values of ln(Elisa) on post-desinewed samples versus the likelihood of positive cnt class, together with a regression line. The derived cutoff value for distinguishing positive samples is the dotted line.



1135
 1136 The establishment effect is clearly seen.

1137
 1138
 1139 **Verification (follow-up) samples**

1140
 1141 From twenty-three establishments, 68 verification samples were collected and analyzed
 1142 for the presence of spinal cord tissue. Most of these samples were collected after the
 1143 survey was completed. These sample results were not included in the analyses presented
 1144 above. One establishment had 19 verification samples, one had 12, and the other
 1145 establishment had no more than 3. Table 10 presents summaries results of the analyses
 1146 matched by establishment, with results from the survey. As can be seen, the percentage of
 1147 positive results for the two establishments with 12 and 19 results are similar, but for the
 1148 remainder, the percentage for the follow-up samples (28%) is less than that of the survey
 1149 samples (37%). It can be seen that the establishments with the smaller fraction of
 1150 positive samples for the survey had generally the smaller fraction of positive verification
 1151 samples.

1152
 1153
 1154
 1155
 1156

1157 Table 10. Summary of fraction of positive results for spinal cord tissue for follow-survey
 1158 up verification samples and survey samples by establishment.

1159
 1160
 1161
 1162
 1163
 1164
 1165

Obs	number samples survey	fraction positive survey	number samples follow-up	fraction positive follow-up
1	6	0.000	1	0.000

1166	2	13	0.000	1	0.000
1167	3	7	0.143	1	0.000
1168	4	6	0.167	1	0.000
1169	5	6	0.167	2	0.000
1170	6	12	0.167	2	0.000
1171	7	10	0.200	3	0.000
1172	8	5	0.200	1	0.000
1173	9	7	0.286	1	0.000
1174	10	6	0.333	19	0.316
1175	11	6	0.333	1	1.000
1176	12	6	0.333	1	0.000
1177	13	6	0.333	1	0.000
1178	14	15	0.333	2	0.000
1179	15	7	0.429	12	0.500
1180	16	6	0.667	1	0.000
1181	17	6	0.667	3	1.000
1182	18	6	0.667	2	0.000
1183	19	6	0.667	2	0.500
1184	20	4	0.750	3	0.667
1185	21	13	0.769	3	0.000
1186	22	6	0.833	3	0.667
1187	23	7	0.857	1	1.000
1188					
1189					

1190 Table 11 presents by the time the samples were analyzed: periods 1 and 2 correspond to
 1191 the first two-thirds and the last third of the survey, and period 3 refers to the follow-up
 1192 samples that were collected after the completion of the survey. The results for the
 1193 establishments with 12 and 19 results are presently separately; for the other
 1194 establishments the results are grouped together and assigned the value of 1 for the
 1195 variable group in Table 11. As seen in Table 11, the percentage of positive samples
 1196 collected during the survey periods are not significantly different, but the percentage of
 1197 the follow-up samples is lower. This pattern holds when considering the establishments'
 1198 percentages of positive samples; Table 12 presents the fractions of positive samples for
 1199 the group 1 samples, defined above, where the samples are further divided by whether or
 1200 not the fraction of sc samples from the survey are greater than 1/3.
 1201

1231 Table 12: Fraction of positive spinal cord (sc) samples by period of time for Group 1
1232 samples, consisting of results from all establishments except the two establishments with
1233 12 and 19 follow-up samples. Samples are classified as to whether or not the sample's
1234 establishment had more then 1/3 of the survey samples with detected spinal cord.

1235
1236 ..ffffffffffffffffffff...fffffffffffffffffffff+
1237 , type
1238 , †ffffffffffff...ffffffffffff%
1239 , follow survey
1240 , †ffffffffffff†ffffffffffff%
1241 , ,frac +, ,frac +,
1242 , N , SC , N , SC ,
1243 †ffffffff...ffffffffff†fff†fffff†fff†fffff%
1244 ,highPos ,period , , , , ,
1245 †ffffffff†ffffffffff%
1246 ,0 ,1 4, 0.000, 68, 0.132,
1247 , †ffffffff†fff†fffff†fff†fffff%
1248 , ,2 1, 0.000, 37, 0.324,
1249 , †ffffffff†fff†fffff†fff†fffff%
1250 , †ffffffff†ffffffffff†fff†fffff†fff†fffff%
1251 ,1 ,1 7, 0.571, 35, 0.657,
1252 , †ffffffff†fff†fffff†fff†fffff%
1253 , ,2 3, 0.667, 19, 0.895,
1254 , †ffffffff†fff†fffff†fff†fffff%
1255 , ,3 8, 0.375,
1256 †ffffffff<ffffffffff†fff†fffff†fff†fffff%
1257 ,All , 36, 0.278, 159, 0.384,
1258 \$ffffffffffffffffffff<fff<fffff<fff<fffff%
1259

1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275

Reference:

Price, Cindy G.; Webb, Neil B.; Smith, Wertice J.; Marks, Harry M.; Yoffe Aaron M.
1994, "Comparison of Mercury and Copper based catalysts in the Kjeldahl determination
of nitrogen in meat and meat products: collaborative study", J. of AOAC International,
vol. 77, 6: p. 1542-1556.

1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296

Attachment:

Derivation of excess iron limits for meat products produced by Advanced Recovery Systems.

1/8/03

1297

1298

**Determining the maximum acceptable level of excess iron
in meat products produced by advanced meat recovery
systems**

1299

1300

1301

1302

1303

Introduction

1304

1305

1306

1307

1308

1309

1310

1311

As discussed in the FSIS response to comments presented in the preamble to Docket 96-027P, FSIS is using an excess iron measurement for evaluating process control because this measure is associated with bone marrow in the product. The assumption is that there is a significant probability that more than negligible amounts of bone marrow would be present in product with elevated excess iron measurements. If an obtained excess iron measurement is larger than a statistically defined amount, then the obtained measurement is considered elevated.

1312

1313

1314

1315

1316

1317

1318

1319

1320

1321

1322

The objective of the rule, stated in the preamble of the 1998 proposal is to "provide clear standards... that include adequate markers for bone-related components (levels consistent with defects anticipated when meat is separated by bone by hand)." This objective is interpreted to mean that for advanced meat recovery (AMR) product to be labeled meat, the excess iron measured levels should be no higher than worst case levels expected (or anticipated) for meat derived from hand deboning when produced under acceptable manufacturing practices. Thus if a product produced by advanced recovery systems has excess iron measured levels greater than these worst case excess iron measured levels, then there is a significant probability that the high iron levels in the product are due to the incorporation of more than negligible bone marrow into the product.

1323

1324

1325

1326

1327

1328

1329

1330

1331

1332

1333

1334

1335

1336

1337

1338

1339

1340

A statistical criterion for determining that a specified product (lot) was produced under acceptable manufacturing practice is derived by considering the distribution of an appropriate product characteristic (such as excess iron) when the product is produced under acceptable manufacturing practices and choosing a percentile, p , of this distribution as a demarcation value, $D(p)$. Thus, if, for product produced in a lot, the measured characteristic is greater than $D(p)$, it is assumed that the lot was not produced under acceptable manufacturing practice. The confidence that this is a true assumption and thus a correct decision is greater than p , or, in other words, there is less than a $(1-p)$ probability that the product actually was produced under acceptable manufacturing practice even though the decision was made that it was not so produced. The choice of p is often based on an assessment of the relative costs and risks associated with incorrect decisions, and, lacking some compelling reason, is often set between 95% to 99.9%. The choice of 99.9% would correspond to approximately 3 standard deviation units when the distribution is symmetric and normal. Using 3 standard deviation units is a common choice in quality control when there is desire to be highly confident that a decision to reject a product as being produced under good manufacturing practice is a correct decision. For the regulation, a choice of 99.9% confidence or 3 standard deviations units above a specified target is used for determining all tolerances.

1341 To determine the distribution of excess iron measurements in hand-deboned meat
1342 product, the measurement error due to repeatability will be accounted for. Information
1343 from USDA's Agriculture Research Service (ARS) is used to establish a repeatability
1344 standard deviation of 0.16 mg/100g for a single iron determination. The data and the
1345 results of statistical analysis of the data are presented as an attachment to this report.
1346 The repeatability standard deviation of protein is set equal to $0.03x^{0.64}$ where x is the %
1347 protein content obtained using the Kjeldahl procedure with mercury catalyst (Price,
1348 Cindy G.; Webb, Neil B.; Smith, Wertice J.; Marks, Harry M.; Yoffe Aron M. 1994,
1349 "Comparison of Mercury and Copper based catalysts in the Kjeldahl determination of
1350 nitrogen in meat and meat products: collaborative study", J. of AOAC International,
1351 vol. 77, 6: p. 1542-1556).
1352

1353 **Maximum mean level (MML) for a lot**

1354

1355 The term "lot" in this setting is used to represent product produced by advanced
1356 recovery systems that has been processed uniformly. It is assumed that the starting
1357 materials used, the calibrations of the machinery, and other processing parameters that
1358 affect the composition of the product would be as uniform as possible. Thus, a lot does
1359 not necessarily represent the product produced in a day. Within a lot, the excess iron
1360 measurements for different samples of the product would be different due to
1361 unavoidable differences in ratios of iron to protein in different animals and analytical
1362 variations in the measured iron and protein levels in samples. However, the lot would
1363 have a mean excess iron level, which would reflect processing control and would
1364 provide, therefore, an appropriate measure for evaluation. In accordance with the
1365 above objective of determining the excess iron measured limits of product produced by
1366 advanced recovery systems that can not be labeled meat, the first step is to determine
1367 the maximum mean level, MML, of excess iron for a lot. From the discussion in the
1368 previous section, the MML is equal to 3 times the "between lot" standard deviation of
1369 the excess iron for meat derived by hand deboning, prior to the bone-in material being
1370 processed by the recovery system. Once this level is determined, then compliance
1371 criteria, based on chemical analysis of samples, are developed which take into
1372 consideration the between sample and analytical measurement variability. In
1373 particular, the criterion for an individual sample, based on duplicate analyses (for both
1374 protein and iron) is derived.

1375 In order to derive the excess iron MML for a "lot" and a criterion for an
1376 individual sample, the 1996 FSIS AMR backbone survey results for the meat derived
1377 from hand deboning will be used. From each of two establishments, 27 samples of
1378 meat derived from hand deboning were collected on various days of production with 3
1379 samples a day (Table 1). The FSIS procedure to measure iron employed a hydraulic
1380 wet acid digestion procedure. However, another method, performed by ARS scientists,
1381 which uses a dry ash procedure for digestion, obtained iron results approximately
1382 double those originally obtained by FSIS. Furthermore, the results obtained by the
1383 ARS dry ash procedure were more consistent with levels reported in the HNS
1384 Handbook 8 levels for hand deboned meat. Consequently, the excess iron values will
1385 be calculated using iron results obtained by the ARS dry ash procedure. For samples
1386 for which there were not ARS dry - ash procedure results, the FSIS results were

1387 multiplied by 2.12 (which was the average ratio of the dry - ash procedure results to the
1388 FSIS results). For the 54 samples of hand-deboned product, 45 of them were analyzed
1389 by the ARS dry - ash procedure. Table 2 provides a comparison of all the FSIS and
1390 ARS obtained results.

1391 In actuality, the meat derived from hand deboning in the survey might have been
1392 heterogeneous, so that within a day there might be more than one "lot" of homogeneous
1393 product. In an analysis of variance, the day within an establishment effect had a
1394 significance level (p- value) of 0.15 (based on 16 degrees of freedom). An
1395 examination of the data did not reveal any particular result or set of results that could be
1396 classified as an outlier. This suggests that the between day variability compared to the
1397 within day variability was not relatively large and that a single day might consist of
1398 more than one lot. Thus, it would be expected that the actual between lot variance
1399 might be larger than the measured between day variance. Since the between sample
1400 (within day) variance is considerably larger than the repeatability variance, it is possible
1401 that a sample "represents" a lot. The "truth" may actually be between the two
1402 extremes, identified here, of a day representing a lot or a sample representing a lot.
1403 Thus a "compromise" calculation is used for determining the between lot variance
1404 component. Specifically, the between lot variance component is set equal to "p"
1405 percent of the within day variance component plus the between day variance
1406 component, and the within lot variance is set equal to $1-p\%/100$ of the measured within
1407 day variance component. For the regulatory derived criteria, p was set equal to 50%,
1408 so that the between lot variance is assumed to equal the sum of the between day
1409 variance plus $\frac{1}{2}$ of the between sample/within day variance, and the within lot variance
1410 is assumed to be equal to $\frac{1}{2}$ the between sample/within day variance.

1411 Excess iron for hand-deboned product, ExFe, is computed as iron minus 0.138
1412 times the percentage protein ($Fe - 0.138\text{protein}$). The factor 0.138 is the ratio of the
1413 average iron to average protein of the hand deboned neckbone product from the FSIS
1414 survey, so that, for this product, the mean of the excess iron results is 0.00 mg/100g.
1415 This factor is also equal to the average iron to protein ratios of the samples. As stated
1416 above, the repeatability standard deviation for the iron measurements is assumed to be
1417 equal to 0.16 mg/100g, which was derived from information obtained from ARS (see
1418 attachment), and the repeatability of protein measurements is equal to $0.03x^{0.64}$ where x
1419 is the percent protein in the sample. Using these values, from the formula for excess
1420 iron, $ExFe = Fe - 0.138x$, the average repeatability variance from the hand deboned
1421 samples was calculated to be equal to 0.0265. An analysis of variance (AOV) of the
1422 sample excess iron results is presented in Table 3.

1423 The between day/establishment variance, from Table 3, is estimated to be
1424 0.13408, and the between sample/within day variance, after accounting for the
1425 measurement variance is estimated to be 0.2875. Thus, the between lot variance,
1426 assuming $p = \frac{1}{2}$, is $0.13408 + \frac{1}{2} 0.2875 = 0.2778$, so that between lot standard
1427 deviation is 0.5271 mg/100g. The maximum mean excess iron for a lot (MML) is 3
1428 times the between lot standard deviation = $3(0.5271) = 1.5813$ mg/100g.

1429

1430

1431

Determining Tolerance for Compliance Purposes

1432

1433 FSIS may take samples to evaluate whether or not establishments are producing
 1434 product produced by advanced recovery systems with “lot” averages greater than the
 1435 MML, 1.5813 mg/100g. The amount of product in a sample is assumed the same as the
 1436 sample amounts of the 1996 FSIS survey of product derived from advanced recovery
 1437 systems.

1438 FSIS recognizes that the ARS process removes connective tissue that contains
 1439 “little or no iron.” FSIS believes the effect of this removal is not large. Connective
 1440 tissues can be removed pre- or post-desinewing. The amount that is removed during
 1441 the pre-desinewing stage of processing depends on the machine pressure applied when
 1442 separating the meat from the bone; the higher the pressure, the more connective tissue
 1443 is removed. From the FSIS 1996 survey, it seems that the average difference in protein
 1444 between pre- and post-desinewing product was 0.5 percent, based on a post-desinewing
 1445 product average protein of about 16.5 percent. Therefore, as a percentage of protein,
 1446 the amount of protein associated with connective tissue removed during this step
 1447 averaged about 3 percent and does not represent a large proportion of the protein that is
 1448 in the product.

1449 In addition, during the ARS processing, some unbound water is removed which
 1450 would result in the removal of some water-soluble protein and dissolved solids. A
 1451 possible consequence therefore is that some water-soluble proteins are removed and
 1452 most of the bound iron will remain in the product, thus, resulting in a higher iron to
 1453 protein ratio in the ARS product.

1454 Because of these two reasons for the possible increase in iron to protein ratio of
 1455 ARS product, for this final rule, FSIS is incorporating 10 percent factor to adjust the
 1456 protein when calculating levels of excessive iron in ARS product. Thus, in the
 1457 calculations of excess iron, the measured protein will be multiplied by 1.10.

1458 A general sampling plan is to take n samples throughout the lot, composite them,
 1459 and perform n_r analytical measurements. If Fe_i and pr_i represent the i^{th} iron and protein
 1460 results, respectively, then the adjusted excess iron, $aExFe$, result for an n – sample
 1461 composite is

1462

$$aExFe = \sum Fe_i / n_r - (0.138)(1.10) \sum pr_i / n_r$$

1463

1464 The expected variance of the adjusted excess iron estimator, $aExFe$, for n – sample
 1465 composites obtain by such a sampling plan is:

$$\text{var}(aExFe) = \sigma_r^2 / n + \sigma_s^2 / n_r$$

1466

1467 where σ_r^2 is the repeatability variance of adjusted excess iron measurements and σ_s^2 is
 1468 the between sample/within lot variance of the adjusted excess iron.

1469 The above identified within lot variance component is determined using the
 1470 results obtained from the 1996 FSIS AMR backbone survey on the hand deboned
 1471 product, using the ARS dry-ash iron results, as explained above. FSIS is using the
 1472 results from this product rather than the AMR product because FSIS considers that the
 1473 AMR product was not produced in accordance with the FSIS requirements for meat,
 1474 and thus, can not, justifiably, be used for determining the within lot standard deviation
 1475 for product produced by advanced recovery systems that is comparable to meat. It

1476 might be that, under good manufacturing practices, product produced by advanced
1477 recovery systems would be more homogeneous than its counterpart meat derived from
1478 hand deboning, so that the within lot variance for such produced AMR product would
1479 be smaller than the within lot variances derived here.

1480 In order to select a specific sampling plan (that is, the number of samples for a
1481 lot) producer and consumer risks (probabilities of the lot passing the test) must be
1482 selected. The MML represents the maximum mean level for a lot that does not result in
1483 a non-compliance determination, thus if a lot had a mean equal to the MML then there
1484 should be a high probability that this lot would not fail and pass the sampling plan. As
1485 discussed above, for determining tolerances, FSIS is selecting 3 standard deviations
1486 above the mean, so that if a result on a n – sample composite is obtained that exceeds
1487 the demarcation value, then there would be approximately 99.9% confidence that the
1488 mean for the “lot” exceeds the MML, and thus the product within the lot would not be
1489 considered comparable to meat. To compute a consumer risk, the probabilities of
1490 passing a lot with mean excess iron level that is equal to 2 times the MML are
1491 determined.

1492 FSIS laboratories would analyze a compliance sample at least in duplicate (see
1493 attachment). Thus it is assumed that n – sample composites are analyzed in duplicate,
1494 so that $n_r = 2$. Because of the factor 1.10, the variance components for this estimator
1495 will be different from those given in Table 3. The analysis of variance for the meat
1496 derived from hand deboning was repeated using the above formula. Presented in Table
1497 4, are the derived variance components for the above estimator of the adjusted excess
1498 iron statistic for the hand deboned product.

1499 The protein values do not affect by much the standard deviation, so that it can be
1500 assumed that the repeatability variance is 0.0267. The between sample/within lot
1501 variance, σ_s^2 , as discussed above, is equal to $(1-p)$ times the between sample/within
1502 day, where $p = \frac{1}{2}$. Thus, $\sigma_s^2 = \frac{1}{2} 0.2905 = 0.1453$, and the expected variance for the
1503 adjusted excess iron results for n – sample composites is therefore, $\text{Var}(a\text{ExFe}) =$
1504 $0.1453/n + 0.0267/2$. The square root of this quantity is the expected standard
1505 deviation.

1506 Table 5 provides demarcation values for determining that a lot has mean excess
1507 iron greater than the MML, 1.5813 mg/100g, for different numbers of samples taken
1508 from the lot, assuming duplicate measurements on the composite of the samples. An
1509 individual sample is when $n=1$ so that the individual sample limit that is specified in the
1510 regulation is 2.776 mg/100g. For purposes of the regulation (for recalling the
1511 demarcation value), this is adjusted to 2.800 mg/100g, so that if an obtained sample
1512 result (based on the average of duplicate analyses of iron and protein) is greater than or
1513 equal to 2.800 mg/100g then the product produced by advanced recovery systems can
1514 not be labeled meat (see conclusion section, below).

1515 If a different percentage, p , than 50% of the within day variance component is
1516 added to the between day variance component, then different answers are obtained for
1517 the individual sample demarcation value and for MML. Figure 1 is a plot of the MML
1518 and the individual sample limit. The percentage that gives the maximum individual
1519 sample demarcation value is 30% and the maximum value is 2.8048 mg/ 100g, with a
1520 MML equal to 1.400 mg/100g. The minimum possible derived individual sample

1521 limit, obtained when $p=100\%$, is 2.2942 mg/ 100g, with the maximum possible derived
1522 MML equal to 1.948 mg/100g.

1523

1524 **FSIS Survey of product produced by advanced recovery systems.**

1525

1526 Presented in Table 6 are the establishment means of adjusted excess iron for
1527 product produced by advanced recovery systems, after the 10% adjustment, $aExFe = Fe$
1528 $-(0.138)(1.10)(\text{protein})$, and the percentage of samples that are greater than or equal to
1529 the derived individual sample limit, 2.800 mg/100g. All the establishment means are
1530 greater than the MML of 1.5813 mg/100g. Also included in Table 6 is the
1531 establishment means of excess iron(not adjusted) of the meat derived from hand
1532 deboning. The highest individual excess iron sample result for the hand-deboned meat
1533 was 1.76 mg/100g. For product produced by advanced recovery systems, 62% of the
1534 samples had adjusted excess iron results that were greater than or equal to 2.800
1535 mg/100g.

1536

1537

1538 **Conclusion:**

1539

1540 If a mean of results from duplicate analyses on a sample is greater than or equal
1541 to 2.800 mg/100g then it is assumed that there is product that is not meat, because of
1542 the incorporation of more than a negligible amount of bone marrow. The question that
1543 needs to be answered is to what product (the lot) does this conclusion apply. In
1544 answering this question, it is assumed that contiguous product is in the same lot. Since
1545 an establishment is required to have documentation that its production process is in
1546 control, it is assumed that a non-compliant finding is a result of a failure or a deficiency
1547 in the process control. A consequence is that all product that is produced before or after
1548 the non-compliant sample might also have been produced when the process was not in
1549 control and thus should or could not be labeled meat. One way of showing that product
1550 is not from the same "lot" is to examine the records of values of processing parameters
1551 that affect the composition of the product produced by advanced recovery systems or
1552 other analytical results from samples of product produced in different parts of the day
1553 or on different days and to determine if there are reasons to identify different "lots"
1554 which would not have non-complying product (mean levels of excess iron less than
1555 1.58 mg/100g).

1556

1557
1558
1559
1560
1561
1562
1563
1564

Figure 1: Plot of derived maximum mean excess iron for lot (MML) and individual sample excess iron limit as function of percentage, where within lot variance equals the sum of the between day variance plus given percentage of within day variance and within lot variance equals 100-percentage of within day variance.

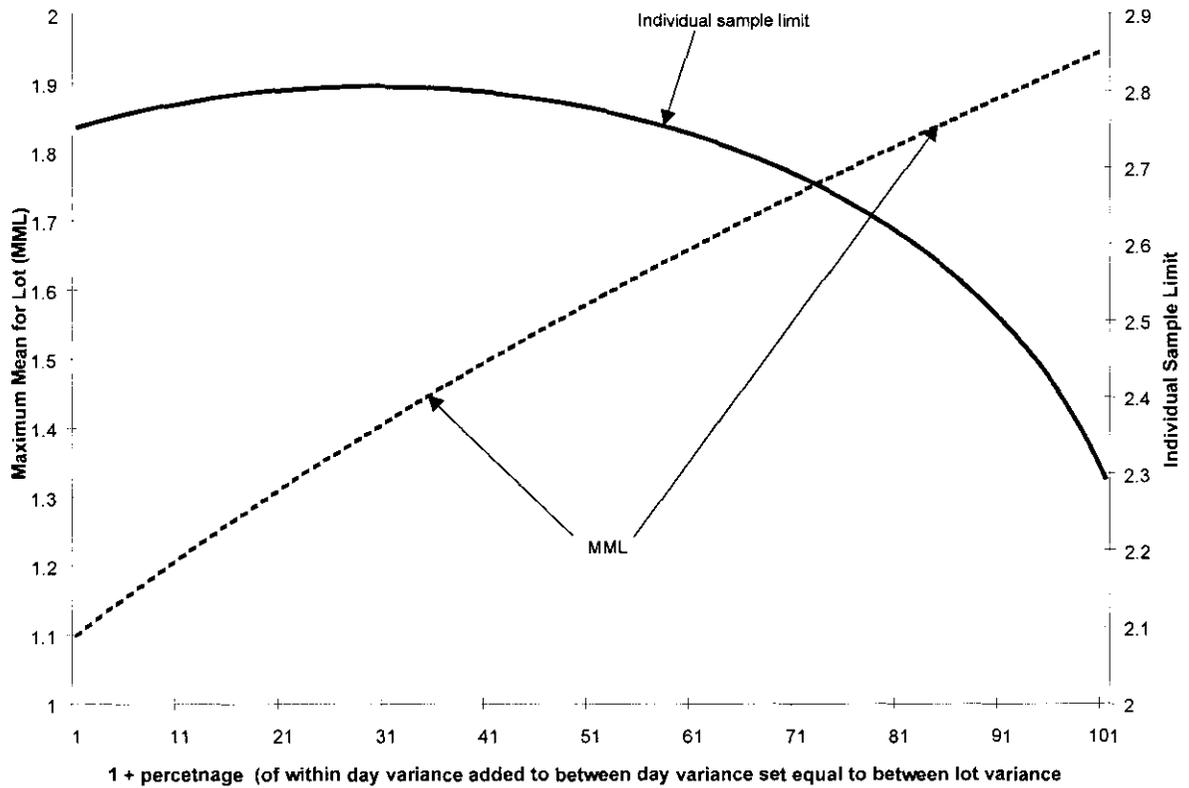


Table 1: Iron and Protein measurements obtained for hand deboned product. obtained from 1996 FSIS survey.

OBS	Est. code	Date of sampling code	Iron dry-ash mg/100g	Protein(%)	Excess iron measure mg/100g
1	5	1	2.26	19.13	-0.38
2	5	1	2.36	19.59	-0.34
3	5	1	2.46	19.32	-0.21
4	5	2	2.74	19.56	0.04
5	5	2	2.60	20.24	-0.19
6	5	2	3.29	19.61	0.59
7	5	3	2.54	21.90	-0.48
8	5	3	2.47	21.38	-0.48
9	5	3	4.01	22.65	0.88
10	5	4	2.50	21.25	-0.43
11	5	4	2.62	21.81	-0.39
12	5	4	2.47	22.72	-0.66
13	5	5	3.04	19.91	0.29
14	5	5	3.22	19.50	0.53
15	5	5	3.79	19.33	1.13
16	5	6	3.22	21.76	0.22
17	5	6	3.44	22.15	0.39
18	5	6	4.14	20.85	1.26
19	5	7	3.22	23.05	0.04
20	5	7	2.99	23.35	-0.23
21	5	7	4.33	21.96	1.30
22	5	8	3.03	23.30	-0.19
23	5	8	2.92	22.95	-0.25
24	5	8	4.55	23.25	1.34
25	5	9	3.18	23.25	-0.03
26	5	9	4.93	23.00	1.76
27	5	9	3.52	22.00	0.48
28	8	1	2.70	20.88	-0.19
29	8	1	2.64	21.54	-0.33
30	8	1	2.43	21.49	-0.54
31	8	2	2.88	21.86	-0.14
32	8	2	2.84	21.07	-0.07
33	8	2	2.88	22.34	-0.21
34	8	3	2.58	21.56	-0.40
35	8	3	2.73	21.91	-0.29
36	8	3	2.62	21.02	-0.28
37	8	4	2.56	21.64	-0.42
38	8	4	2.72	22.92	-0.45
39	8	4	2.48	22.14	-0.57
40	8	5	3.01	22.80	-0.14
41	8	5	3.71	21.09	0.80

Table 1 (cont): Iron and Protein measurements obtained for hand deboned product.

OBS	Est. code	Date of sampling code	Iron dry-ash mg/100g	Protein(%)	Excess iron measure mg/100g
42	8	5	3.35	21.44	0.39
43	8	6	2.45	22.02	-0.59
44	8	6	2.14	20.75	-0.72
45	8	6	1.67	21.79	-1.34
46	8	7	2.98	20.90	0.09
47	8	7	2.31	23.91	-0.99
48	8	7	3.66	22.10	0.61
49	8	8	3.35	22.90	0.19
50	8	8	3.23	23.20	0.03
51	8	8	3.40	22.61	0.28
52	8	9	2.37	20.42	-0.45
53	8	9	1.94	21.84	-1.07
54	8	9	3.82	21.91	0.80

1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645

Table 2: Comparison of ARS Dry Ash and FSIS Wet Acid digestion results units mg/100g (AMR= product from advanced recovery systems)

OBS	Type of Product	FSIS Wet Acid	ARS Dry Ash	Ratio Dry Ash to Wet Acid
1	AMR	2.36	5.08	2.15
2	AMR	3.19	5.90	1.85
3	AMR	2.46	7.03	2.86
4	AMR	1.83	4.13	2.26
5	AMR	2.69	4.13	1.54
6	AMR	2.81	5.15	1.83
7	AMR	2.49	4.80	1.93
8	AMR	1.79	5.18	2.89
9	AMR	2.23	5.59	2.51
10	AMR	2.41	5.97	2.48
11	AMR	7.91	8.32	1.05
12	AMR	4.88	7.02	1.44
13	AMR	2.39	5.56	2.33
14	AMR	2.94	5.23	1.78
15	AMR	2.57	4.99	1.94
16	AMR	2.64	5.08	1.92
17	AMR	2.72	4.97	1.83
18	AMR	2.81	4.88	1.74
19	AMR	2.82	5.09	1.80
20	AMR	2.04	5.26	2.58
21	AMR	3.59	5.36	1.49
22	AMR	3.56	5.19	1.46
23	AMR	2.90	6.17	2.13
24	AMR	3.03	5.43	1.79
25	AMR	2.52	3.53	1.40
26	AMR	3.51	5.61	1.60
27	AMR	3.26	5.33	1.63
28	AMR	3.32	5.03	1.52
29	AMR	2.68	4.76	1.78
30	AMR	3.17	5.42	1.71
31	AMR	3.90	6.05	1.55
32	AMR	2.31	5.00	2.16
33	AMR	2.70	6.43	2.38
34	AMR	1.51	4.93	3.26
35	AMR	2.36	5.37	2.28
36	AMR	2.26	5.07	2.24
37	AMR	1.79	5.26	2.94
38	AMR	1.70	4.44	2.61
39	AMR	2.42	4.37	1.81
40	AMR	1.70	4.89	2.88
41	AMR	2.30	6.20	2.70
42	AMR	2.51	6.31	2.51
43	AMR	2.52	5.61	2.23
44	AMR	2.66	6.17	2.32
45	AMR	3.05	4.72	1.55
46	AMR	3.21	6.11	1.90
47	AMR	2.39	6.10	2.55

Table 2 (cont): Comparison of ARS Dry Ash and FSIS Wet Acid digestion results
Units mg/100g (AMR= product from advanced recovery systems)

	OBS	Type of Product	FSIS Wet Acid	ARS Dry Ash	Ratio Dry Ash to Wet Acid
700					
701					
702					
703					
704					
705					
706					
707					
708	48	AMR	2.20	4.99	2.27
709	49	AMR	1.78	6.11	3.43
710	50	AMR	2.20	5.85	2.66
711	51	AMR	2.42	5.91	2.44
712	52	AMR	2.55	6.38	2.50
713	53	AMR	3.87	6.78	1.75
714	54	AMR	1.78	4.86	2.73
715	55	AMR	1.79	6.74	3.77
716	56	AMR	2.27	5.95	2.62
717	57	AMR	4.11	6.43	1.56
718	58	AMR	2.96	6.19	2.09
719	59	AMR	2.31	4.99	2.16
720	60	AMR	2.37	4.72	1.99
721	61	AMR	2.29	4.67	2.04
722	62	AMR	1.94	5.42	2.79
723	63	AMR	1.52	5.43	3.57
724	64	AMR	2.07	5.62	2.71
725	65	AMR	3.03	5.63	1.86
726	66	AMR	2.05	7.47	3.64
727	67	AMR	2.36	6.58	2.79
728	68	AMR	3.00	5.30	1.77
729	69	AMR	2.89	4.85	1.68
730	70	AMR	3.21	4.36	1.36
731	71	AMR	2.84	5.44	1.92
732	72	AMR	2.69	5.89	2.19
733	73	AMR	2.95	6.09	2.06
734	74	AMR	3.34	6.33	1.90
735	75	AMR	3.95	5.91	1.50
736	76	AMR	4.44	7.52	1.69
737	77	AMR	3.45	5.37	1.56
738	78	AMR	3.73	5.47	1.47
739	79	AMR	3.69	5.94	1.61
740	80	AMR	2.75	4.89	1.78
741	81	AMR	2.51	5.69	2.27
742	82	AMR	2.62	5.65	2.16
743	83	AMR	2.69	5.93	2.20
744	84	AMR	2.36	5.07	2.15
745	85	AMR	1.97	4.42	2.24
746	86	AMR	2.02	5.40	2.67
747	87	AMR	2.25	5.91	2.63
748	88	AMR	3.63	8.95	2.47
749	89	AMR	3.85	6.98	1.81
750	90	AMR	4.08	7.13	1.75
751	91	AMR	2.57	5.70	2.22
752	92	AMR	3.03	8.19	2.70
753	93	AMR	1.88	3.26	1.73
754	94	AMR	2.64	6.52	2.47

Table 2(cont): Comparison of ARS Dry Ash and FSIS wet Acid digestion results
Units mg/100g (AMR= product from advanced recovery systems)

OBS	Type of Product	FSIS Wet Acid	ARS Dry Ash	Ratio Dry Ash to Wet Acid	
755					
756					
757					
758					
759					
760					
761					
762					
763					
764	95	AMR	2.54	7.27	2.86
765	96	AMR	3.86	7.01	1.82
766	97	AMR	3.10	6.39	2.06
767	98	AMR	3.77	5.71	1.51
768	99	AMR	3.02	6.80	2.25
769	100	AMR	2.35	5.16	2.20
770	101	AMR	2.58	4.86	1.88
771	102	AMR	2.29	5.02	2.19
772	103	AMR	2.96	3.87	1.31
773	104	AMR	3.21	5.91	1.84
774	105	AMR	1.88	5.93	3.15
775	106	AMR	2.04	5.48	2.69
776	107	AMR	2.40	4.14	1.73
777	108	AMR	2.36	5.11	2.17
778	109	AMR	3.10	4.99	1.61
779	110	AMR	3.42	5.83	1.70
780	111	AMR	2.34	4.86	2.08
781	112	AMR	3.89	4.91	1.26
782	113	AMR	2.96	6.05	2.04
783	114	AMR	3.50	6.43	1.84
784	115	AMR	3.37	5.38	1.60
785	116	AMR	1.97	4.96	2.52
786	117	AMR	3.15	5.60	1.78
787	118	AMR	3.00	6.15	2.05
788	119	AMR	3.24	5.96	1.84
789	120	AMR	2.72	5.68	2.09
790	121	AMR	3.54	6.00	1.69
791	122	AMR	3.70	6.55	1.77
792	123	AMR	2.63	4.83	1.84
793	124	AMR	3.12	5.95	1.91
794	125	AMR	4.14	8.21	1.98
795	126	AMR	1.78	4.37	2.46
796	127	AMR	1.96	5.27	2.69
797	128	AMR	1.91	3.98	2.08
798	129	AMR	2.22	5.84	2.63
799	130	AMR	2.85	5.35	1.88
800	131	AMR	2.35	6.13	2.61
801	132	AMR	3.64	7.67	2.11
802	133	AMR	3.19	7.08	2.22
803	134	AMR	2.27	6.56	2.89
804	135	AMR	2.96	5.16	1.74
805	136	AMR	2.57	6.16	2.40
806	137	AMR	2.25	6.00	2.67
807	138	AMR	2.14	6.37	2.98
808	139	AMR	2.39	6.61	2.77
809	140	AMR	2.81	6.66	2.37
	141	AMR	5.30	6.34	1.20

Table 2(cont): Comparison of ARS Dry Ash and FSIS Wet Acid digestion results
Units mg/100g (AMR= product from advanced recovery systems)

OBS	Type of Product	FSIS Wet Acid	ARS Dry Ash	Ratio Dry Ash to Wet Acid
142	AMR	5.13	6.98	1.36
143	AMR	4.36	5.13	1.18
144	Hand	1.06	2.26	2.13
145	Hand	1.10	2.64	2.40
146	Hand	1.09	2.36	2.17
147	Hand	1.06	2.46	2.32
148	Hand	1.34	2.43	1.81
149	Hand	1.40	2.70	1.93
150	Hand	1.78	3.29	1.85
151	Hand	1.56	2.60	1.67
152	Hand	1.82	2.74	1.50
153	Hand	1.78	2.88	1.62
154	Hand	1.90	2.88	1.51
155	Hand	1.79	2.84	1.58
156	Hand	1.43	2.48	1.74
157	Hand	1.56	2.62	1.68
158	Hand	1.65	2.72	1.65
159	Hand	1.48	2.50	1.69
160	Hand	1.37	2.56	1.87
161	Hand	1.60	2.62	1.64
162	Hand	1.18	2.47	2.09
163	Hand	1.26	2.58	2.05
164	Hand	1.69	1.67	0.99
165	Hand	1.53	2.45	1.60
166	Hand	1.21	2.14	1.77
167	Hand	1.74	3.04	1.75
168	Hand	1.70	3.82	2.25
169	Hand	1.07	1.94	1.81
170	Hand	1.51	2.37	1.57
171	Hand	1.01	3.01	2.98
172	Hand	1.25	2.47	1.98
173	Hand	1.20	4.01	3.34
174	Hand	1.14	2.54	2.23
175	Hand	1.41	3.22	2.28
176	Hand	1.77	4.14	2.34
177	Hand	1.64	3.44	2.10
178	Hand	1.14	2.98	2.61
179	Hand	1.58	3.66	2.31
180	Hand	1.62	2.31	1.43
181	Hand	1.19	3.23	2.71
182	Hand	1.10	3.40	3.09
183	Hand	1.19	3.35	2.82
184	Hand	1.22	3.03	2.48
185	Hand	1.60	4.55	2.84
186	Hand	1.44	2.92	2.03
187	Hand	1.46	4.93	3.38
188	Hand	1.40	4.33	3.09

1865
 1866
 1867
 1868
 1869
 1870
 1871

Table 3: Analysis of variance of excess iron results, ExFe, from 1996 FSIS survey of hand- deboned neckbone samples (2 establishments with 27 observations per establishment). $\text{ExFe} = \text{iron} - 0.138\text{protein}$

Source of Variation	Variance	Standard deviation
Between establishment	0.08054	0.2838
Between Day within establishment	0.05354	0.2314
Sum: Between establishment/day	0.13408	0.3662
Within day including measurement error	0.3140	0.5603
Measurement error	0.0265	0.1628
Between sample/Within day	0.2875	0.5362
Total variance	0.4481	0.6694

1872
 1873
 1874
 1875
 1876
 1877
 1878
 1879
 1880
 1881

Table 4: Analysis of variance of adjusted excess iron, aExFe, based on results from 1996 FSIS survey of meat derived from hand deboning. $\text{aExFe} = \text{iron} - (0.138)(1.10)\text{protein}$.

Source of Variation	Variance	Standard Deviation
Between establishment/day	0.1333	0.3651
Within day including measurement error	0.3172	0.5632
error	0.0267	0.1633
Between sample/Within day	0.2905	0.5390

1882
 1883

1884

1885 Table 5: Limits for determining that a lot has mean adjusted excess iron, aExFe, greater
 1886 than 1.58mg/100g for n – sample composites as function of the number of samples, n,
 1887 per lot, assuming duplicate analysis on the composite of the samples. The aExFe n –
 1888 sample composite result is equal to the mean of the iron results minus the product of
 1889 0.138, 1.10 and the mean protein result. The derived limit is equal to 3 expected
 1890 standard deviations above the maximum mean for a lot (MML) = 1.5813 mg/100g.
 1891 Also presented are the probabilities of passing a lot with a true excess iron mean =
 1892 3.1626 mg/100g (= 6 between lot standard deviations above zero excess iron).

1893

Number of Samples	Limit	Prob. (%) passing lot mean=3.163
1	2.776	16.5821
2	2.461	0.8345
3	2.327	0.0385
4	2.250	0.0021
5	2.199	0.0001

1894

1895

1896

1897

1898 Table 6: Summary of excess iron results from 1996 FSIS neckbone survey.

1899 The hand-deboned excess iron results are computed as: iron – 0.138protein,
 1900 The product produced by advanced recovery systems (AMRS) excess iron results
 1901 are computed as: iron – (0.138)(1.10)protein.

1902

establish- ment code	number of samples	mean excess iron	percent samples > 2.776
8 hand	27	0.221	0.00
9 hand	27	-0.221	0.00
all hand	54	0.000	0.00
1 ^a AMRS	27	2.778	40.74
2 AMRS	24	3.950	87.50
3 AMRS	16	3.280	56.25
4 AMRS	27	2.656	33.33
5 AMRS	25	3.443	76.00
6 ^b AMRS	25	3.560	84.00
7 AMRS	19	3.065	57.90
ALL AMRS	163	3.235	61.96

1917

1918 ^a) establishment used Protecon machine, while others used Hydrosep
 1919 machines.

1920

1921

1922

1923

^b) establishment did not perform desinewing operation.

1924
1925
1926
1927

Attachment:

Repeatability of iron measurements using the ARS Dry-Ash procedure

1928 Data to determine the repeatability of the ARS Dry- Ash procedure was provided
1929 to FSIS by Dr. Bob Windham of ARS. Analyses were conducted on beef samples. For
1930 further details contact Dr. Bob Windham. The first data set consists of duplicate results
1931 obtained by the same laboratory on 47 samples. The second data set are results from a 3-
1932 laboratory, 5-sample collaborative study, where each sample was analyzed in duplicate
1933 by each lab.

1934 The results from the 47 samples of the first data set are given in Table 1.
1935 Statistical analysis did not indicate a non-zero correlation of the standard deviations and
1936 mean levels of the samples, so that it is assumed that the repeatability standard deviation
1937 does not depend upon the level of iron in the sample. Figure 1 is a plot of the sample
1938 standard deviations versus the sample means for the 47 samples. The line represents a
1939 quadratic fit. It can be seen from this graph the 5 data points that have standard
1940 deviations greater than 0.5 mg/ 100g. The standard deviations of these 5 data points can
1941 be assumed to be outlier standard deviations. This can be seen by computing the ratio of
1942 the maximum sample variance to the sum of the sample variances and comparing this
1943 ratio to appropriate percentiles of a beta distribution (Hawkins, D. M., 1980,
1944 Identification of Outliers, Chapman and Hall, New York, NY, Appendix 9).
1945 Specifically, let $v_{(j)}$ be a random variable representing the j^{th} ordered sample variance
1946 from k samples. The ratio of the maximum sample variance to the sum of the sample
1947 variances,

$$r_k = v_{(k)} / \sum_{j=1}^k v_{(j)}$$

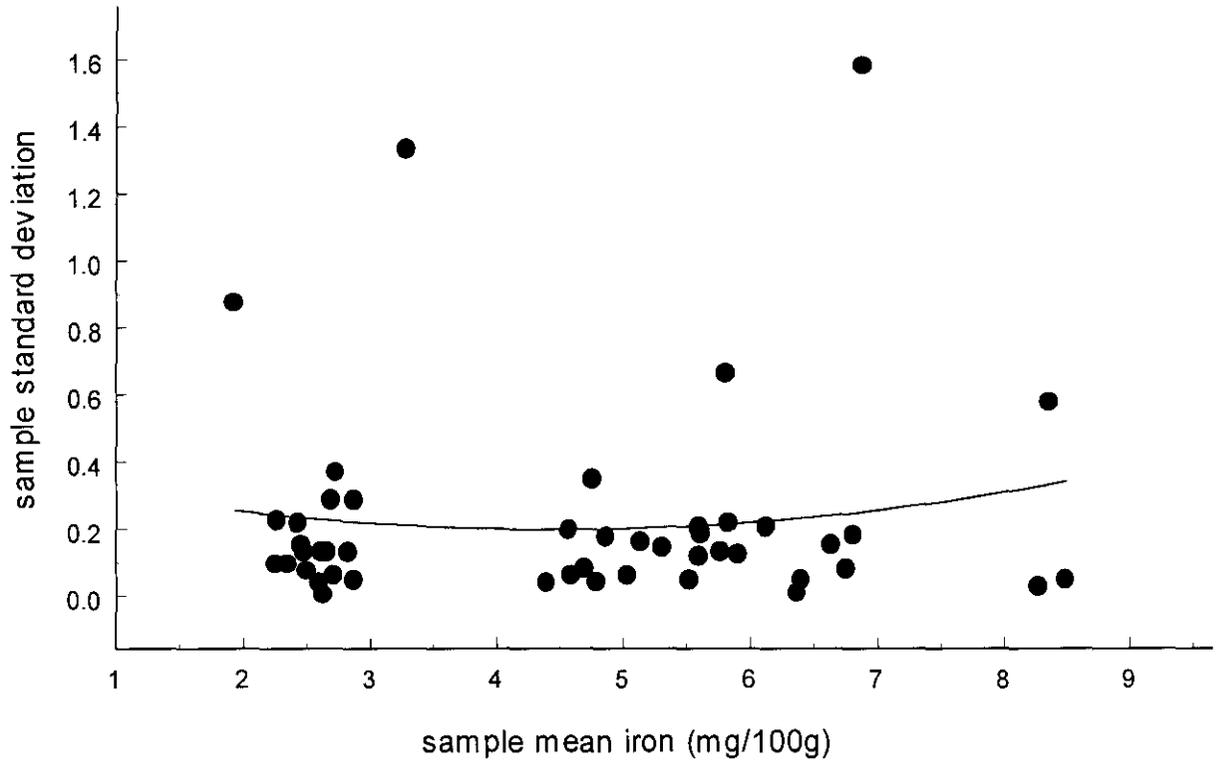
1948 is compared to an appropriate percentile of a beta distribution with parameters $\frac{1}{2}$ and $(k-$
1949 $1)/2$. To determine whether $v_{(k)}$ is an outlier with respect to the set $\{v_{(j)}\}$, for $j < k$, the
1950 observed value of the ratio, r_k , is compared to the $1 - \alpha/k$ percentile of the beta distribution,
1951 where α represents the significance of the statistical test of $v_{(k)}$ being an outlier. For
1952 $k=43, \dots, 47$, the ratios r_k were computed and the corresponding significance levels, α_k ,
1953 were determined. For $k=47$, $\alpha_{47} = 0.00029$, so that the highest computed variance can be
1954 considered as an outlier. For $k=46$, $\alpha_{46} = 0.00007$, so that the second highest variance
1955 can be considered as an outlier. Also, $\alpha_{45} = 0.00482$, $\alpha_{44} = 0.03097$ and $\alpha_{43} = 0.03577$,
1956 so that the five highest variances can be considered as outliers. Assuming that the
1957 variances on these five samples are outlier results and thus excluding them from the
1958 analysis, the repeatability standard deviation from the remaining 42 samples is estimated
1959 (by computing the square root of the mean of sample variances) to be 0.161 mg/100g. If
1960 the sample with standard deviation 0.58 mg/100g is included in the calculations, then the
1961 estimated repeatability standard deviation is estimated to be 0.182 mg/100g. Further, the
1962 distribution of the differences of the duplicate analyses within the 42 samples appeared to
1963 be normally distributed. Thus, percentiles of the measurement distribution can be
1964 assumed to be distributed as normal.
1965

1966 The data (Table 2) from the collaborative study (3 labs, 5 samples, measured in
1967 duplicate) contained possible two outlier results. The 4.91 mg/100g result obtained by
1968 the second lab for the first replicate of the third sample is quite different from the other
1969 five results, which range from approximately 8 to 9 mg/100g. Thus, this result was not
1970 used in the statistical analysis. In addition, for the fifth sample, the sample standard
1971 deviation obtained by the first lab, also appears to be an outlier. This can be seen by
1972 examining Table 3, which presents means and standard deviations of the replicate results
1973 for a sample. The computed ratio, r_{14} , of the maximum sample variance to the sum of the
1974 14 sample variances (excluding the third sample from the second lab) is 0.723, which has
1975 statistical significance of $p = 0.0008$. The estimated standard deviation of repeatability
1976 (obtained through an analysis of variance), excluding only the outlier result of 4.91
1977 mg/100g was 0.182 mg/100g. When the results for the fifth sample that were obtained by
1978 the first lab are also deleted, the estimated standard deviation of repeatability is 0.100
1979 mg/100g.

1980 For deriving the criteria for excess iron in AMR product that can be labeled meat,
1981 the repeatability standard deviation is assumed to be 0.16 mg/100g. This is based on the
1982 estimated repeatability standard deviation obtained when deleting the 5 samples with
1983 standard deviations greater than or equal to 0.58 mg/100g. Support for the 0.16 mg/100g
1984 value is the 0.10 mg/100g estimate of the repeatability standard deviation from the
1985 collaborative study when the two outlier results are deleted. Because of the few large
1986 differences of duplicate sample results, it is recommended there should be at least
1987 duplicate analyses on samples used for compliance purposes.

1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998

Figure 1: Plot of within sample standard deviations versus sample mean iron level.



1999 Table 1: Duplicate iron results (mg/100g) from 47 meat samples

sample	replicate		mean	standard deviation	
	1	2			
2000	1	2.33	2.19	2.26	0.10
2001	2	2.63	2.64	2.64	0.01
2002	3	2.29	2.43	2.36	0.10
2003	4	2.35	2.57	2.46	0.16
2004	5	2.58	2.27	2.43	0.22
2005	6	2.90	2.49	2.70	0.29
2006	7	4.24	2.35	3.30	1.34
2007	8	2.63	2.57	2.60	0.04
2008	9	2.47	3.00	2.74	0.37
2009	10	2.84	2.91	2.88	0.05
2010	11	2.67	3.08	2.88	0.29
2011	12	2.74	2.93	2.84	0.13
2012	13	2.58	2.39	2.49	0.13
2013	14	2.72	2.53	2.63	0.13
2014	15	2.67	2.76	2.72	0.06
2015	16	2.56	2.45	2.51	0.08
2016	17	6.95	6.69	6.82	0.18
2017	18	4.64	4.76	4.70	0.08
2018	19	4.52	5.02	4.77	0.35
2019	20	4.83	4.77	4.80	0.04
2020	21	4.74	4.99	4.87	0.18
2021	22	4.37	4.43	4.40	0.04
2022	23	5.82	6.00	5.91	0.13
2023	24	5.42	5.21	5.32	0.15
2024	25	4.55	4.64	4.60	0.06
2025	26	5.67	5.98	5.83	0.22
2026	27	5.56	5.49	5.53	0.05
2027	28	5.52	5.69	5.61	0.12
2028	29	6.38	6.37	6.38	0.01
2029	30	6.70	6.82	6.76	0.08
2030	31	5.03	5.26	5.15	0.16
2031	32	5.68	5.87	5.78	0.13
2032	33	6.44	6.37	6.41	0.05
2033	34	8.30	8.26	8.28	0.03
2034	35	8.46	8.53	8.50	0.05
2035	36	7.95	8.77	8.36	0.58
2036	37	4.71	4.43	4.57	0.20
2037	38	6.53	6.75	6.64	0.16
2038	39	6.29	5.35	5.82	0.66
2039	40	5.45	5.74	5.60	0.21
2040	41	5.48	5.74	5.61	0.18
2041	42	5.00	5.09	5.05	0.06
2042	43	5.77	8.01	6.89	1.58
2043	44	6.27	5.98	6.13	0.21
2044	45	2.10	2.42	2.26	0.23
2045	46	2.76	2.57	2.67	0.13
2046	47	2.55	1.31	1.93	0.88

2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067

Table 2: Results from Collaborative study; data provided by ARS

lab	repli- cation	Sample number				
		1	2	3	4	5
1	1	5.68	6.44	8.30	8.46	8.77
1	2	5.87	6.37	8.26	8.53	7.95
2	1	4.96	5.92	4.91	7.41	7.90
2	2	4.73	5.96	8.00	7.25	8.05
3	1	5.81	5.21	8.99	8.77	8.88
3	2	5.79	5.53	9.03	8.73	8.84

Table 3: Means and standard deviations for samples

SAMPLE	Laboratory									
	1		2		3		pool			
	MEAN	STD	MEAN	STD	MEAN	STD	MEAN	STD		
1	5.775	0.134	4.845	0.163	5.800	0.014	5.473	0.122		
2	6.405	0.049	5.940	0.028	5.370	0.226	5.905	0.135		
3	8.280	0.028	8.000		9.010	0.028	8.516	0.028		
4	8.495	0.049	7.330	0.113	8.750	0.028	8.192	0.073		
5	8.360	0.580	7.975	0.106	8.860	0.028	8.398	0.341		