***Description of study herds***

Of the 21 herds included, 5 used a bedded pack system (23.8%), 1 used a freestall bedded with sand (4.8%), 5 used a freestall bedded with shavings/sawdust (23.8%), and 10 used a tiestall bedded with shavings/sawdust (47.6%) (Supplemental Table S1). The 21 farms were spread out over 9 of Vermont’s 14 counties: Orleans and Franklin County (4 each), Orange (5), Addison and Chittenden (2 each), and Caledonia, Lamoille, Washington, and Windham (1 each). The predominant breeds for each farm were Holstein (8), Jersey (10), and mixed Holstein-Jersey crosses/other (3). The median (mean; range) number of lactating cows was 68 (64.9; 32-99). The median rolling herd average for the farms was 14,037 (14,163; 9,000-21,204) pounds. Nineteen of the 21 farms tested with DHIA monthly while their cows were in milk, with one farm testing 5-8 times/year and one testing every other month. Detailed descriptions further characterizing study farm management practices and housing characteristics for lactating animals (e.g., laying surface, ventilation, stocking density), and details about bedding material and bedding management practices for lactating animals (e.g., bedding depth, frequency of adding new bedding, manure removal) are provided in Supplemental Tables S1 and S2, respectively. Detailed descriptions of routine milking procedures and mastitis control practices are provided in Supplemental Tables S3 and S4, respectively.

***Comparison of bulk tank milk aerobic culture data***

Of the 21 bulk tank milk samples cultured, none were positive for *Strep. agalactiae* or *Mycoplasma* spp. (Table 1). Sixteen of the 21 bulk tank milk samples were negative for coliforms on aerobic culture, while 5 farms had a coliform count of 5 cfu/mL. *Staph. aureus* was found in the bulk tank milk from 13/21 herds, with a median (mean; range) cfu/mL of 50 (70; 15-320) when present. The median (mean; range) *Staph.* spp. count found in the 21 bulk tank milk samples was 65 (96; 0-665) cfu/mL, while the median non-*ag. Strep.* count was 45 (156; 10-1250) cfu/mL. As most measures of the aerobic culture data were not normally distributed even after log transformation, a Kruskal-Wallis test was used to compare CFU counts between the three facility types. There was no difference in mean rank in respect to the three facility types for CFU counts of *Staph.* spp., non-*ag Strep.* spp., coliforms, or *Staph. aureus* (p > 0.05, Table 1).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1. Bulk tank milk aerobic culture outcomes by facility type for 21 Vermont organic dairy herds. | | | | | | | | |
|  |  | Count |  |  | Mean | 95% CI | Range | Kruskal-Wallis χ2 |
| *Staph.* spp. (CFU/mL) | | 21 |  |  | 95.5 | 36-155 | 0-665 | 0.9 (df = 2, p = 0.62) |
|  | Bedded pack | 5 |  |  | 53 | 10-96 | 0-130 |  |
|  | Tiestall | 10 |  |  | 134.5 | 14-255 | 15-665 |  |
|  | Freestall | 6 |  |  | 65.9 | 28-104 | 5-125 |  |
|  |  |  |  |  |  |  |  |  |
| Non-*ag. Strep.* (CFU/mL) | | 21 |  |  | 156.4 | 42-271 | 10-1250 | 4.6 (df = 2, p = 0.10) |
|  | Bedded pack | 5 |  |  | 39 | 17-61 | 10-80 |  |
|  | Tiestall | 10 |  |  | 255.5 | 30-481 | 20-1250 |  |
|  | Freestall | 6 |  |  | 89.2 | 11-167 | 25-260 |  |
|  |  |  | Neg. for *S. aureus* (no. farms) | Pos. for *S. aureus* (no. farms) |  |  |  |  |
| *Staph. aureus* (CFU/mL) | | 21 | 8 | 13 | 43.6 | 14-73 | 0-320 | 3.4 (df = 2, p = 0.19) |
|  | Bedded pack | 5 | 3 | 2 | 9 | 0-21 | 0-30 |  |
|  | Tiestall | 10 | 3 | 7 | 63 | 5-121 | 0-320 |  |
|  | Freestall | 6 | 2 | 4 | 40 | 10-70 | 0-100 |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  | Neg. for coliforms(no. farms) | 5 CFU/mL coliforms  (no. farms) | Mean | 95% CI | Range |  |
| Coliforms (CFU/mL) | | 21 | 16 | 5 | 1.2 | 0.3-2.1 | 0-5 | 0.4 (df = 2, p = 0.82) |
|  | Bedded pack | 5 | 4 | 1 | 1 | 0-3 | 0-5 |  |
|  | Tiestall | 10 | 7 | 3 | 1.5 | 0.003-3 | 0-5 |  |
|  | Freestall | 6 | 5 | 1 | 0.8 | 0-2.5 | 0-5 |  |
|  |  |  |  |  |  |  |  |  |
| *Strep. agalactiae* (CFU/mL) | | | all 21 farms negative | |  |  |  |  |
| *Mycoplasma* spp. (CFU/mL) | | | all 21 farms negative | |  |  |  |  |

***Comparison of bulk tank milk udder health measures by facility type***

The mean (SD; range) raw somatic cell count for the 21 bulk tank milk samples collected was 144,286 cells/mL (53,934; 54,000-250,000) (Table 2). For the 19 herds with available DHIA test-day data, the mean percent of cows with new IMI was 5.7 (3.7; 0-12.3), the mean percent of cows with chronic IMI was 13.6 (5.8; 2.9-23.1), and the mean percent of cows with any IMI was 25 (7.8; 8.6-36.9). For the 18 herds with available data, the mean standardized 150-day milk was 50 pounds (10.1; 33.5-68). For the 20 herds with available data, the mean average unweighted linear score of cows was 2.44 (0.42; 1.7-3.3), while the mean average linear score weighted for production was 3.72 (0.56; 2.4-5).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table 2. Udder health and production outcomes by facility type for 21 Vermont organic dairy herds. | | | | | | |
|  |  |  | Continuous outcomes | | |  |
| Outcome |  | Sample size | Mean | 95% CI | Range |  |
| Bulk tank milk somatic cell count (cells/mL) | | 21 | 144,286 | 121,218-167,353 | 54,000-250,000 | One-way Fisher’s ANOVA: (F(2,18) = [2.137], p = 0.15) |
|  | Bedded pack | 5 | 107,600 | 70,275-144,925 | 54,000-160,000 |  |
|  | Tiestall | 10 | 146,400 | 114,136-178,664 | 97,000-250,000 |  |
|  | Freestall | 6 | 171,333 | 126,965-215,702 | 98,000-250,000 |  |
|  |  |  |  |  |  |  |
| Percent of cows with new IMI since last test date (%) | | 19 | 5.7 | 4.2-7.3 | 0-12.3 | One-way Fisher’s ANOVA: (F(2,16) = [0.2128], p = 0.81) |
|  | Bedded pack1 | 3 | 7.0 | 2.8-11.2 | 2.9-12.3 |  |
|  | Tiestall | 10 | 5.4 | 3.0-7.8 | 0-10.5 |  |
|  | Freestall | 6 | 5.6 | 3.0-8.3 | 0-9.8 |  |
|  |  |  |  |  |  |  |
| Percent of cows with chronic IMI since last test date (%) | | 19 | 13.6 | 11.2-16.1 | 2.9-23.1 | One-way Fisher’s ANOVA: (F(2,16) = [0.3138], p = 0.74) |
|  | Bedded pack1 | 3 | 14.5 | 5.4-23.7 | 2.9-23.1 |  |
|  | Tiestall | 10 | 14.3 | 11.9-16.7 | 7.8-20.8 |  |
|  | Freestall | 6 | 12.0 | 6.7-17.3 | 5.7-23.1 |  |
|  |  |  |  |  |  |  |
| Percent of cows with SCC ≥ 200,000 cells/mL on current test date (%) | | 19 | 24.9 | 21.6-28.3 | 8.6-36.9 | One-way Fisher’s ANOVA: (F(2,16) = [0.1017], p = 0.90) |
|  | Bedded pack1 | 3 | 26.0 | 12.6-39.3 | 8.6-36.9 |  |
|  | Tiestall | 10 | 25.4 | 22.1-28.6 | 17.6-32.8 |  |
|  | Freestall | 6 | 23.7 | 16.9-30.5 | 11.6-36.5 |  |
|  |  |  |  |  |  |  |
| Standardized 150-day milk (pounds) | | 18 | 50.0 | 45.7-54.3 | 33.5-68.0 | One-way Fisher’s ANOVA: (F(2,15) = [0.4404], p = 0.65) |
|  | Bedded pack2 | 4 | 46.9 | 39.8-53.9 | 38.5-56.3 |  |
|  | Tiestall2 | 8 | 49.4 | 43.1-55.7 | 33.5-68.0 |  |
|  | Freestall | 6 | 53.0 | 43.5-62.5 | 38.7-67.7 |  |
|  |  |  |  |  |  |  |
| Avg. linear score of cows on farm (unweighted) | | 20 | 2.44 | 2.26-2.62 | 1.7-3.3 | One-way Welch’s ANOVA: (F(2,5.7) = [0.02891], p = 0.97) |
|  | Bedded pack2 | 4 | 2.38 | 1.84-2.91 | 1.7-3.1 |  |
|  | Tiestall | 10 | 2.45 | 2.31-2.59 | 2.2-2.8 |  |
|  | Freestall | 6 | 2.5 | 2.00-2.93 | 1.9-3.3 |  |
|  |  |  |  |  |  |  |
| Avg. linear score of cows on farm (weighted by production) | | 20 | 3.72 | 3.48-3.96 | 2.4-5.0 | Kruskal-Wallis  χ2 = 0.32, df = 2,  p = 0.85 |
|  | Bedded pack2 | 4 | 3.48 | 2.77-4.18 | 2.4-4.3 |  |
|  | Tiestall | 10 | 3.75 | 3.52-3.98 | 3.4-4.4 |  |
|  | Freestall | 6 | 3.83 | 3.27-4.40 | 3.0-5.0 |  |
| 1 DHIA data not available for 2 farms | | | | | |  |
| 2 DHIA data not available for 1 farm | | | | | |  |

***Comparison of hygiene scores by facility type***

The overall mean (SD; range) of herd-level hygiene scores for all 21 farms was 2.32 (0.39; 1.7-3.0). The mean (SD; range) hygiene score was 2.2 (0.31; 1.7-2.48) for bedded pack farms (n = 5), 2.5 (0.42; 1.9-3.04) for tiestall farms (n = 10), and 2.15 (0.28; 1.8-2.78) for freestall farms (n = 6). A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference in mean hygiene score of cows in a herd between any of the three facility types (F(2,18) = [2.307], p = 0.13).

The overall mean (SD; range) proportion of cows with dirty udders in a herd (udder hygiene score ≥ 3) was 40% (20; 10-79%). The proportion of cows with dirty udders (SD; range) was 32% (16; 10-48%) for bedded pack farms (n = 5), 49% (21; 21-79%) for tiestall farms (n = 10), and 32% (16; 16-55%) for freestall farms (n = 6). A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference proportion of cows with dirty udders in a herd between any of the three facility types (F(2,18) = [2.1031], p = 0.15).

***OR:***

The overall mean (95% CI; range) of herd-level hygiene scores for all 21 farms was 2.32 (2.16-2.49; 1.7-3.0). The mean (95% CI; range) hygiene score was 2.2 (1.91-2.44; 1.7-2.48) for bedded pack farms (n = 5), 2.5 (2.24-2.76; 1.9-3.04) for tiestall farms (n = 10), and 2.15 (1.93-2.37; 1.8-2.78) for freestall farms (n = 6). A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference in mean hygiene score of cows in a herd between any of the three facility types (F(2,18) = [2.307], p = 0.13).

The overall mean (95% CI; range) proportion of cows with dirty udders in a herd (udder hygiene score ≥ 3) was 40% (31-48%; 10-79%). The proportion of cows with dirty udders (95% CI; range) was 32% (18-46%; 10-48%) for bedded pack farms (n = 5), 49% (35-62%; 21-79%) for tiestall farms (n = 10), and 32% (20-44%; 16-55%) for freestall farms (n = 6). A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference proportion of cows with dirty udders in a herd between any of the three facility types (F(2,18) = [2.1031], p = 0.15).

***Results of univariate analyses for udder health and hygiene outcomes***

Initially, we had intended on using a modelling approach (linear regression) to explore each of the six udder health and hygiene outcomes by facility type, while controlling for different farm management practices, housing characteristics, milking procedures and mastitis control practices. In preparation for this, an initial univariable analysis was performed for each outcome to screen all predictors to check for unconditional associations at a level of p ≤ 0.20 (using linear regression). While this model-building approach to analysis was abandoned due to sample size/statistical power limitations, the results for each of the six outcome variables are reported in Table 8 (“biologically sensical,” and groups of at least n = 5).

The depth of bedding in stalls (cm) for freestall and tiestall herds was unconditionally associated with multiple udder health outcomes, including new IMI, any IMI, BTM SCC, and both weighted and unweighted average LS. As the depth of bedding in stalls increased, these udder health measures improved (lower LS, BTM SCC, percent of any/new IMI). Similarly, herds where cows were on deep bedding had a lower BTM SCC compared to herds that had stalls with a smaller amount of bedding on top of a mattress or concrete. Udder hygiene measures were associated with numerous udder health outcomes, including chronic IMI, any IMI, and both weighted and unweighted average LS. A higher proportion of dirty cows (increasing average hygiene score of a farm, higher proportion of udders scored ≥ 3) was associated with higher percentages of chronic IMI, any IMI, and both weighted and unweighted average LS. As these two udder hygiene measures were both calculated from the same data for each herd, they are highly correlated with one another, and conclusions were not drawn distinguishing the effect of one vs. another for each udder health outcome. Additionally, a few specific management practices were also found to be unconditionally associated with udder health outcomes; consistent glove was associated with a lower percentage of new IMI, clipping or flaming udders was associated with fewer chronic IMI, and both parenteral supplementation of a Vitamin E/selenium product and an approved intramammary product at dry-off were associated with a lower unweighted average LS.

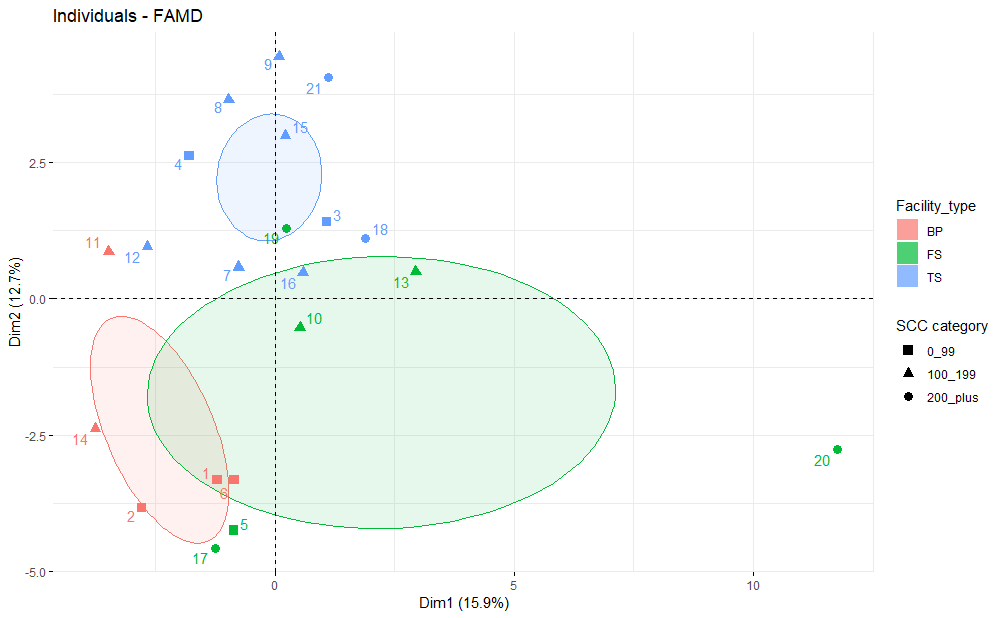
Both udder hygiene outcomes were unconditionally associated with the same predictors, all of which were related to the depth of bedding for cows. For herds using a bedded pack, the deeper the bedding was at the time of measurement the cleaner the udders were. For cows in tiestalls and freestalls, cleaner udders were associated increasing depth of bedding in stalls. For all herds, cows on deep bedding (bedded packs or deeply-bedded freestalls) had better udder hygiene than cows on a smaller amount of bedding on top of a mattress or concrete.

***OR:***

Farms with deeper bedding showed a tendency toward a lower bulk tank SCC (deep bedding lying surface vs. mattress/concrete, p = 0.14; depth of bedding in stalls, p = 0.06), lower % new IMI (depth of bedding in stalls, p = 0.02), lower % any IMI (p = 0.01), lower weighted and unweighted average LS (depth of bedding in stalls p = 0.05 and 0.10, respectively), lower mean udder hygiene score (deeply-bedded stalls vs. mattress/concrete, p = 0.06; depth of bedding in stalls p = 0.07), and lower proportion of dirty udders (deeply-bedded stalls vs. mattress/concrete, p = 0.06; depth of bedding in stalls p = 0.13). Farms with lower mean udder hygiene scores tended towards having lower % chronic IMI (proportion dirty udders and mean hygiene p = 0.05), lower % any IMI (proportion dirty udders, p = 0.13; mean hygiene, p = 0.09), and lower weighted (proportion dirty udders, p = 0.20) and unweighted average LS (proportion dirty udders, p = 0.12; mean hygiene, p = 0.11). Increased bedding depth measures also tended to be associated with lower mean udder hygiene scores (deeply-bedded stalls vs. mattress/concrete, p = 0.06; depth of bedding in stalls, p = 0.07; bedded pack depth p = 0.01), as well as lower proportion of dirty udders (deeply-bedded stalls vs. mattress/concrete, p = 0.06; depth of bedding in stalls, p = 0.13; bedded pack depth p ≤ 0.001).

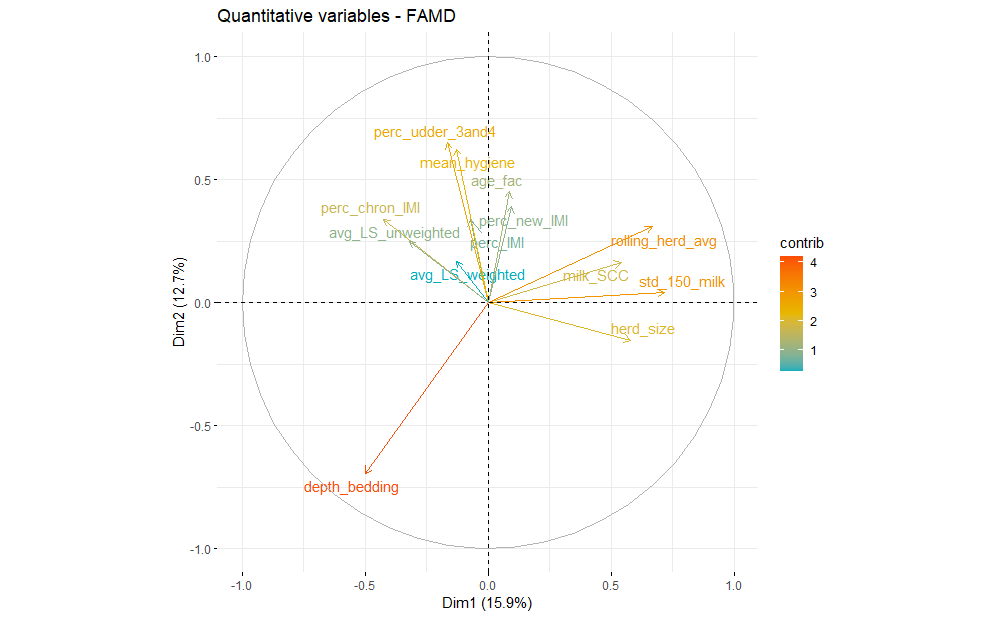
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table 3. Selected results of univariate analysis identifying factors unconditionally associated with udder health and hygiene outcomes at p < 0.20 for 21 Vermont organic dairy herds | | | | |
| Parameter | | Group (sample size) | Coefficient estimate (SE) | *P-*value |
| Bulk tank milk somatic cell count (cells/mL) | |  |  |  |
|  | Lying surface | Mattress or concrete (n = 13) | 36,000 (23,454) | 0.14 |
|  |  | Deep bedding (n = 8) | Ref. | Ref. |
|  | Depth of bedding in stalls (cm)1 | Tiestalls and freestalls (n = 15) | -5,797 (3,970) | 0.17 |
| % cows with new IMI since last test date (%)2 | |  |  |  |
|  | Glove use | Inconsistent glove use while milking (n = 9) | 2.83% (1.7) | 0.11 |
|  |  | All milkers consistently use gloves (n = 9) | Ref. | Ref. |
|  | Depth of bedding in stalls (cm)1 | Tiestalls and freestalls (n = 15) | -0.62% (0.24) | 0.02 |
| % cows with chronic IMI since last test date (%)2 | |  |  |  |
|  | Herds that clip or flame udders one or more times per lactation | Yes (n = 5) | -4.31% (2.9) | 0.16 |
|  |  | No (n = 14) | Ref. | Ref. |
|  | Proportion of cows with udder hygiene scores ≥ 3 | All herds with available test data (n = 19) | 12.7% (6) | 0.05 |
|  | Average hygiene score | All herds with available test data (n = 19) | 6.39% (3.1) | 0.05 |
| % cows with SCC ≥ 200,000 cells/mL on current test date (%)2 | |  |  |  |
|  | Depth of bedding in stalls (cm)1 | Tiestalls and freestalls (n = 15) | -1.2% (0.42) | 0.01 |
|  | Proportion of cows with udder hygiene scores ≥ 3 | All herds with available test data (n = 19) | 13.6% (8.5) | 0.13 |
|  | Average hygiene score | All herds with available test data (n = 19) | 7.7% (4.3) | 0.09 |
| Average linear score of cows on farm (unweighted)3 | |  |  |  |
|  | Herds that regularly do parenteral supplementation of dry cows with vit. E and selenium | Regular or occasional supplementation (n = 9) | -0.27 (0.18) | 0.15 |
|  |  | No parenteral supplementation (n = 11) | Ref. | Ref. |
|  | Herds that use any sort of approved organic intramammary product at dry-off | Yes (n = 5) | -0.29 (0.21) | 0.18 |
|  |  | No (n = 15) | Ref. | Ref. |
|  | Depth of bedding in stalls (cm)1 | Tiestalls and freestalls (n = 15) | -0.05 (0.03) | 0.1 |
|  | Proportion of cows with udder hygiene scores ≥ 3 | All herds with available test data (n = 20) | 0.75 (0.45) | 0.12 |
|  | Average hygiene score | All herds with available test data (n = 20) | 0.39 (0.23) | 0.11 |
| Average linear score of cows on farm (weighted)3 | |  |  |  |
|  | Depth of bedding in stalls (cm)1 | Tiestalls and freestalls (n = 15) | -0.06 (0.03) | 0.05 |
|  | Proportion of cows with udder hygiene scores ≥ 3 | All herds with available test data (n = 20) | 0.85 (0.63) | 0.2 |
| Proportion of cows with udder hygiene scores ≥ 3 | |  |  |  |
|  | Depth of bedded pack (m) | Bedded pack herds (n = 5) | -0.5 (0.06) | <0.01 |
|  | Lying surface | Mattress or concrete (n = 13) | 0.17 (0.08) | 0.06 |
|  |  | Deep bedding (n = 8) | Ref. | Ref. |
|  | Depth of bedding in stalls (cm)1 | Tiestalls and freestalls (n = 15) | -0.02 (0.02) | 0.13 |
| Average hygiene score | |  |  |  |
|  | Depth of bedded pack (m) | Bedded pack herds (n = 5) | -0.96 (0.15) | < 0.01 |
|  | Lying surface | Mattress or concrete (n = 13) | 0.33 (0.16) | 0.06 |
|  |  | Deep bedding (n = 8) | Ref. | Ref. |
|  | Depth of bedding in stalls (cm)1 | Tiestalls and freestalls (n = 15) | -0.06 (0.03) | 0.07 |
| 1 Stall bedding depth for freestalls and tiestalls bedded with wood shavings or sawdust | | | | |
| 2 DHIA data available for n = 19 herds. One herd included in average linear score analyses is seasonal and had no recent test data. | | | | |
| 3 DHIA data available for n = 20 herds. | | | | |

***Results of factor analysis of mixed data (FAMD)***



Farms clustered pretty closely by housing system; Bedded pack farms are more similar to each other than they are to other housing system groups (likewise true for tiestalls and freestalls)

Bedded pack farms are capable of achieving excellent milk quality: 3/5 bedded pack farms in the 0-99,000 cells/mL BTSCC category, and the remaining 2 are in the 100,000-199,000 cells/mL BTSCC category (6 farms are 0-99,000, 10 farms are 100-199, 5 farms are 200 plus)



***Comparison of bulk tank milk udder health measures, aerobic culture data, and hygiene scores by facility type***

****

**Figure 1.** Bulk tank milk somatic cell count by facility type. The central line of each box represents the median, while the upper and lower lines represent the upper (75th) and lower (25th) quartiles, respectively. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. A one-way ANOVA revealed that there was not a statistically significant difference in bulk tank somatic cell count between any of the three facility types (F(2,18) = [2.137], p = 0.15).

**OR**

**Figure 1.** Bulk tank milk somatic cell count by facility type. Boxplot showing the 25th, 50th (median), and 75th percentiles of bulk tank raw somatic cell count data from 21 herds. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference in bulk tank somatic cell count between any of the three facility types (F(2,18) = [2.137], p = 0.14).

****

**Figure 2.** Proportion of cows in a herd with a new IMI by facility type, where a new IMI was defined as the linear score changing from <4.0 to ≥4.0 in the last 2 tests. Boxplots showing the 25th, 50th (median), and 75th percentiles of DHIA data available from 19 herds. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. Outliers are points located beyond the upper and lower whiskers. A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference in proportion of cows with a new IMI between any of the three facility types (F(2,16) = [0.2128], p = 0.81).



**Figure 3.** Proportion of cows in a herd with a chronic IMI by facility type, where a chronic IMI was defined as a linear score ≥4.0 on the last 2 tests. Boxplots showing the 25th, 50th (median), and 75th percentiles of DHIA data available from 19 herds. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference in proportion of cows with a chronic IMI between any of the three facility types (F(2,16) = [0.3138], p = 0.74).



**Figure 4.** Proportion of cows in a herd with any IMI by facility type, where any IMI was defined as having a linear score ≥4.0 on most recent test day. Boxplots showing the 25th, 50th (median), and 75th percentiles of DHIA data available from 19 herds. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. Outliers are points located beyond the upper and lower whiskers. A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference in proportion of cows with a new IMI between any of the three facility types (F(2,16) = [0.1017], p = 0.90).



**Figure 5.** Standardized 150 Day Milk by facility type. Boxplots showing the 25th, 50th (median), and 75th percentiles of DHIA data available from 18 herds. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. Outliers are points located beyond the upper and lower whiskers. A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference in Standardized 150 Day Milk between any of the three facility types (F(2,15) = [0.4404], p = 0.65).



*The variances are not homogenous for this analysis…. Violates an assumption of the regular ANOVA; can’t do Kruskal-Wallis with unequal variance either; so did Welch’s ANOVA (normally distributed, but unequal variances)*

**Figure 6.** Average unweighted linear score of cows in a herd by facility type. Boxplots showing the 25th, 50th (median), and 75th percentiles of DHIA data available from 20 herds. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. A Welch’s ANOVA (unequal variances) revealed that there was not a statistically significant difference in average unweighted linear score of cows in a herd between any of the three facility types (F(2,5.7) = [0.02891], p = 0.97).



*Data for tiestall not normally distributed… Violates an assumption of the regular ANOVA, and can’t do Welch’s if not normally distributed; instead, do Kruskal-Wallis (not normally distributed, but does have homogenous variances)*

**Figure 7.** Average linear score of cows weighted by production in a herd by facility type. Boxplots showing the 25th, 50th (median), and 75th percentiles of DHIA data available from 20 herds. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. A Kruskal-Wallis chi-squared test (unequal variances between groups) showed there was no difference in mean rank in respect to the three facility types for average linear score of cows weighted by production (Kruskal-Wallis χ2 = 0.32, df = 2, p = 0.85).



**Figure 8.** Mean hygiene score of cows in a herd by facility type. Boxplots showing the 25th, 50th (median), and 75th percentiles of data from 21 herds. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference in mean hygiene score of cows in a herd between any of the three facility types (F(2,18) = [2.307], p = 0.13).



**Figure 9.** Proportion of cows with dirty udders (udder hygiene score ≥ 3) in a herd by facility type. Boxplots showing the 25th, 50th (median), and 75th percentiles of data from 21 herds. The upper whisker represents the largest observation less than or equal to the 75th quartile plus 1.5 times the interquartile range, while the lower whisker represents the smallest observation greater than or equal to the 25th quartile minus 1.5 times the interquartile range. A one-way Fisher’s ANOVA revealed that there was not a statistically significant difference proportion of cows with dirty udders in a herd between any of the three facility types (F(2,18) = [2.1031], p = 0.15).