***Description of Study Herds***

Of the intended 40 herds to be recruited in the study (10 each from each of 4 housing/bedding combination categories: freestalls with sand, freestalls with wood shavings/sawdust, bedded packs, and tiestalls with wood), 21 herds that were contacted agreed and the survey and sampling were completed throughout April-May, 2019. All herds sampled during this period were housing their cows as they would in the winter months. Completion of the survey and sampling was suspended in mid-May, as farms began turning their cows out to pasture, with the intention of resuming in Spring 2020 to complete the remaining 19 herds. However, due to the COVID-19 pandemic, the decision was made to not resume the survey and sampling, and the final analysis included herds only completed in 2019 (n = 21).

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%) (Table 1). 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. 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 5). 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). Of the 21 bulk tank milk samples cultured, none were positive for *Strep. agalactiae* or *Mycoplasma* spp. (Table 6). 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. The mean (SD; range) of the herd-level hygiene scores was 2.32 (0.39; 1.7-3.0), while the mean of the proportion cows with dirty udders in a herd (udder hygiene score ≥ 3) was 40% (20%; 10-79%) (Table 7). 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 Tables 1 and 2, respectively. Detailed descriptions of routine milking procedures and mastitis control practices are provided in Tables 3 and 4, respectively.

***Comparison of Bulk Tank Milk Udder Health Measures, Aerobic Culture Data, and Hygiene Scores by 3 Facility Types***

****

**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. Boxplots 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 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 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 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 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 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 ANOVA*

**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 one-way ANOVA 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(2,17) = [0.05734], p = 0.94).



*Data for tiestall not normally distributed… Violates an assumption of the ANOVA*

**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 one-way ANOVA revealed that there was not a statistically significant difference in average linear score of cows weighted by production in a herd between any of the three facility types (F(2,17) = [0.4842], p = 0.62).



**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 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 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).

A one-way ANOVA was performed to compare the effect of [independent variable] on [dependent variable].

A one-way ANOVA revealed that there [was or was not] a statistically significant difference in [dependent variable] between at least two groups (F(between groups df, within groups df) = [F-value], p = [p-value]).

Tukey’s HSD Test for multiple comparisons found that the mean value of [dependent variable] was significantly different between [group name] and [group name] (p = [p-value], 95% C.I. = [lower, upper]).

There was no statistically significant difference between [group name] and [group name] (p=[p-value]).

***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<0.2 (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 XX (“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 Avg. 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 Avg. 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 Avg. 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 management practices were found to be unconditionally associated with udder health outcomes as well; 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 Avg. 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.

**NEW IMI**

Gloves

Depth bedding stalls

**CHRONIC IMI**

Clip/flame udders

Prop. Dirty cows

Avg. hygiene score (although correlated with 3+4 cows obvi)

**ANY IMI**

Depth bedding stalls

Prop. Dirty cows

Avg. hygiene score (although correlated with 3+4 cows obvi)

**BTM SCC**

Lying surface

Depth bedding stalls

**AVG LS (UNWEIGHTED)**

Vit E selenium

Dry-off product

Depth bedding stalls

Prop. Dirty cows

Avg. hygiene score (although correlated with 3+4 cows obvi)

**AVG LS (WEIGHTED)**

Depth bedding stalls

Prop. Dirty cows

**PROP 3+4**

Depth bedded pack

Lying surface

Depth bedding stalls

**AVG HYGIENE**

Depth bedded pack

Lying surface

Depth bedding stalls