***The reviewers identified sample size as a major limiting factor of the study. Completeness of the data was rated as fair or poor by the three reviewers. For the paper to move forward, it will be important to clearly address the reviewers’ comments and explain, given the study limitations, what the results can and cannot demonstrate. One reviewer even suggests that the work represents more of a case study than a research paper. Finally, please note that one of the reviewer's comments are included as an attachment.***

AU: Dr. Middleton, thank you for the opportunity to revise our manuscript in response to the reviewer comments. We agree the work has limitations due to sample size constraints. We identified this issue in the original manuscript draft and have extended this discussion at lines xxx. We appreciate the comment about verging on a case study and have revised our inferential statements based on what the observed results can and cannot demonstrate. As a convenience sample, we are careful to note this is a non-probability study design with the associated limitations in making inferences to broader herd populations. Lines 722-735 also address limitations of convenience sampling (selection bias), and that it is inappropriate to make inference from the study population to the target population. The external validity is somewhat limited and we leave it to the reader to make any inferences. We are hesitant to describe this as a case study as we detail in our specific response to reviewer x. Despite these limitations, we appreciate that the reviewers appear to agree there is some merit in reporting our observations. We have copied Reviewer 1 comments from the annotated pdf document so that you can follow our response to those comments.

**Reviewer(s)' Comments to Author:**

We want to thank the reviewers for their careful, evaluative reading of our manuscript, and the accompanying suggestions, requested clarifications, and helpful comments. The process of incorporating the suggested changes has certainly improved our work and made this a stronger manuscript.

***Reviewer: 1***

***Making conclusions from small studies is difficult. While I'm very happy to avoid the 'tyranny' of p-values, if you are going to conclude that there is 'no difference' you have to look at the range of results that are compatible with your data. You don't and you need to.***

AU: Thank you for this comment. If we understand what the reviewer is requesting, we had not included an overall assessment of the size of the estimate effects and their standard errors for each outcome model, as well as an interpretation of possible biological significance of the range of results found. We have now added to this to the Discussion section (lines 617-638), as well as more language acknowledging the limitations of the study (small group size), and the need for future studies to further explore these relationships. An overall summary of these results has also been added to the Conclusion (lines 823-831).

***Further comments from annotated PDF, Reviewer 1***

AU: We have copied those comments as written referencing the original line number of the annotated pdf

***Line 39: Is n of 5 sufficient to make meaningful claims?***

AU: Thank you for this important question. We share the concern that a sample of 5 farms is small, and as we described our initial intent was to enroll 10 farms. This sample size was based on prior estimates of the size of the target population. We have added language that more clearly explains the original study design considerations and the associated logistical constraints we experienced implementing the study, as well as acknowledging the limitations of drawing conclusive results from small sample size (lines 198-226, and 737-754). We believe n of 5 is sufficient to report our findings, as we are also attempting to be transparent and forthright with the limitations of the study. Additional language was added throughout the manuscript acknowledging that while we found no different between the housing types, more research enrolling a larger number of farms is needed to test the hypothesis that milk quality, udder health, and udder hygiene on bedded packs are comparable to other facility types (line 67, line 535, line 791, line 798).

***Line 60: While I am perfectly happy with identifying differences of biological importance rather than statistical ones, you can't just replace a focus on one point value (0) with a focus on another (study difference) - you need to answer the question "did your study rule out biologically important negative effects"?***

AU: If we understand what the reviewer is requesting, we did not rule out biologically important effects, either positive or negative. We have revised the language to be clear that we cannot reject our null hypothesis of no difference between the housing systems, recognizing the variation among the herds in bedding management. We did not have sufficient data to support an alternative hypothesis that bedded packs had worse measures (i.e., higher hygiene scores, higher BTM SCC, higher proportions of cows with increased SCS); in fact we observed numerically better values for BP compared to the two other housing systems, which we find intriguing and leads us to conclude we have sampled herds that are able to achieve good milk quality using the bedded pack systems being recommended in the Northeastern US.

***Line 80: would be useful to properly define what you mean by this term***

AU: Thank you. We recognize that the term “bedded pack” is vague, as it encompasses multiple styles of management and appears to have variable use in the literature and lay-publications. For the purpose of this study, we use the general term “bedded pack” to include both “conventional bedded packs” and “compost bedded packs,” as defined in Bewley et. al (2017). Because a number of authors suggest “conventional bedded packs” are synonymous with straw yards, we have added language to explain the difference between bedded packs and straw yards. To clarify this for the reader, we have defined what we mean by the term “bedded pack” from line 82-86, and further delineate this grouping when describing the enrolled herds (Materials and Methods, lines 185-208). We hope this clarifies our use of the term, and the diversity of management styles under this umbrella currently being used by dairy producers (especially organic farms) in the Northeast US.

***Line 106: is this wast you are referring to by bedded packs [sic]***

AU: see previous comment.

***Line 181: how did this relate to the ten mentioned above***

AU: We are a bit unsure of what the reviewer is asking here. As mentioned in line 162, we aimed to enroll 40 farms for the current study, with 10 farms from each of the four housing/bedding categories described on line 153-156. The number of farms we were able to sample before interruption was 21 herds total: 1 freestall bedded with sand, 5 freestalls bedded with wood shavings/sawdust, 10 tiestalls bedded with wood shavings/sawdust, and 5 bedded packs (as described on lines 179-181). While we intended to include 10 herds of each of the four types of housing/bedding categories, we were unable to resume the study as planned in 2020 due to the COVID 19 pandemic. We hope this explanation provides the information they were looking for.

***Line 297: As these were counts why dd you not consider Poisson or negative binomial [sic]***

AU: Thank you for this suggestion; we also recognized that these were count data, and attempted exploring them using a number of different methods (Poisson, zero-inflated Poisson, negative binomial, and zero-inflated negative binomial, as appropriate). This is referred to on lines 398-401. However, upon reflection after reading this comment, this sentence seems to be more appropriate in Materials and Methods, and has been moved to there (lines 336-339).

***Line 350: So you have three very different systems across 5 farms compounding the difficulties of interpreting the data***

AU: Thank you for this comment, and we agree that this point should be more overtly addressed in the manuscript. In order to achieve this, we have made clear that multiple management styles are grouped under the “bedded pack” term in the Abstract (lines 39-43), defined what we mean by the term “bedded pack” in the Introduction from line 82-97, further delineate this grouping when describing the enrolled herds (Materials and Methods, lines 185-208), as well as added language around this issue to the Discussion (added lines 754-771). We recognize that for this manuscript, the “bedded pack” group represents both farms using a deep bedded pack or “conventional bedded packs” and “compost bedded-packs,” as defined in Bewley et. al (2017), with the modification to distinguish the recent deep bedded pack designs being adopted in the Northeast US from classical straw yards. The smaller than anticipated number of dairy farms using bedded packs for housing lactating dairy cattle in our state created a challenge for enrolling ten herds in our observational study, and we recognize this as a limitation (lines 759-763). However, this diversity is a reflection of how the target population (small-medium, pasture-based organic dairy farms) are actually using this system in the Northeastern U.S. (Benson, 2012).

Although there is a diversity of management styles under this umbrella term, these systems do still share some important things in common: both are enclosed, loose-housing facilities deeply bedded with organic material, in which bedding and waste accumulate throughout the 6–8-month period of time when cows are housed on it and which is only removed once a year. Both systems use carbon-rich substrates on which urine and manure are not removed when bedding material is renewed, which is in contrast to other housing systems. In both types of bedded packs, cows urinate and defecate on the same surface used for laying, and cows’ teats are exposed to bacteria from manure and the organic bedding which presumably acts as a growth medium. Both systems address the desire to provide a loose-housing option where the initial investment of the building is more readily affordable for those with outdated facilities, provide cows with a large, comfortable laying area which allows them to move freely, allows for improved foot and leg health, and allows farms to avoid the need to store and spread large amounts of liquid manure. Both systems provide dairies with a valuable source of compost as a nutrient amendment, through the accumulation of waste and bedding material in the pack, and both are perceived to integrate well into pasture-based farm systems. All of these aspects of these systems are described in the literature (as referenced in the introduction and discussion); our study is novel in that we describe milk quality attributes for these systems.

Recent previous work has primarily focused on describing bedded packs that are actively managed for aerobic composting (Leso et al., 2020). Leso et al. contrasted composting bedded packs managed with daily cultivation with conventional static bedded packs, such as straw yards, noting the reduced cow cleanliness and increased risk of mastitis associated with the latter. While we did not define “straw yards” in our paper, the modern “traditional” or “deep bedded pack” being implemented as winter housing in pasture-based systems in the Northeastern U.S. contrasts to the traditional straw yard, where is often recommended that bedding is completely removed at approximately monthly intervals (for example as described in Shepley et al., 2020; Whistance et al., 2007, Agriculture and Horticulture Development Board 2012. and https://ahdb.org.uk/knowledge-library/loose-yard-management-to-control-environmental-mastitis-in-dairy-cows x et al. date). In Vermont, bedded packs are primarily used for lactating cows by herds that graze dairy cattle from May-November. In these grass-based organic farms, bedding material accumulates over this 6–8-month period where the lactating cattle are housed, until the facility is cleaned out (usually once a year). The bedding material in these structures accumulates to heights of 4 to 5 feet (120 to 150 cm). This is in contrast to straw yard systems, which are typically used for confinement housing cattle year-round. Straw yards have a long tradition of use in the dairy industry, and as such they are hard to narrowly define. A thorough description is provided in Shepley et al. (2020), which is overall in agreement with how the authors perceive this housing system: “The strawyard (SY) treatment provided … concrete flooring topped with rubber mats. … The lying area was bedded with straw at a depth of 20−25 cm, with daily cleaning and addition of straw to the bedding area to maintain bedding cleanliness and depth.” (Shepley 2020), and also in the Agriculture and Horticulture Development Board 2012 publication at <https://ahdb.org.uk/knowledge-library/loose-yard-management-to-control-environmental-mastitis-in-dairy-cows>

While there exists a substantial body of recent work exploring udder health on compost bedded-pack systems, we were unable to find much peer-reviewed literature describing udder health on deep bedded pack systems (in contrast to considerable work on udder health in straw yards: Astiz et. al, 2014; Fregonesi and Leaver, 2001; Fregonesi and Leaver, 2002; Ward et. al 2002; Peeler et al. 2000). While bedded pack systems are not common for housing lactating cows in Vermont, both composting and static systems are used (Andrews et al., 2021). In service to starting to establish a body of work describing udder health on bedded packs used on organic dairies in Vermont, the current work attempts to shed light on a broader spectrum of options used within this loose-housing system. Our current study shows that farms can achieve excellent milk quality using either an untilled, deep bedded pack system or an aerobically composting bedded pack system for indoor housing.

***Line 398: Again did you try approaches for count data***

AU: see previous comment.

***Line 450: So you doidn;'t use a threshold of <0.05 despite stating that you did earlier [sic]***

AU: For Objective 2 (all facility types grouped together, univariate analysis), we stated in the Materials and Methods that “unconditional relationships between the eight outcome variables and independent predictors are reported for a significance level of P ≤0.20” (previously line 338, now 379). In line 492, we attempt to be transparent that we are not stating that all results presented met the threshold for statistical significance.

In an attempt to be clearer about what thresholds were used for statistical significance for Objective 1 (multivariable models, facility type forced in), additional clarification has been added to lines 370-371.

***Table 4: Similar issues across whoi;le of this table. I have only noted some of them [sic]***

AU: See previous comment.

***Neither of these results rule out a biologically meaningful deleterious relationship (referring to BTSCC coefficient estimates by facility type)***

AU: see previous comment

***Did you use a linear regression for this? It would seem to be a candidate for logistic regression as this a proportion affected [sic] (referring to % newly elevated SCS and % chronically elevated SCS outcome models results)***

AU: Thank you for this comment, and we appreciate the suggestion. We have taken this recommendation into consideration, and consulted with a few other epidemiologists who agree that using a linear regression for modelling these outcomes (percent newly-elevated SCS, percent chronically-elevated SCS, percent of cows with SCS ≤4) is an acceptable method of analysis. Although it is true that conceptually the prevalence of these three udder health measures is bounded by 0-100%, there are theoretically an unlimited number of decimal points between 0 and 100 that these values could take. These are herd-level values, taken from DHIA records, for one particular point in time relative to our farm visits. If we were looking to estimate the likelihood of one cow in a herd becoming a case for a new intramammary infection from one test to the next, we agree that a logistic regression would be the appropriate analysis. Although we appreciate this comment, we would like to point out there are other examples of previously-published work in JDS and other peer-reviewed journals using this approach for the same kind of herd-level udder health data:

Patel, K., S. M. Godden, E. Royster, B. A. Crooker, J. Timmerman, and L. Fox. 2019. Relationships among bedding materials, bedding bacteria counts, udder hygiene, milk quality, and udder health in US dairy herds. J. Dairy Sci. 102(11):10213-10234.

In the above study, linear regression was used to model *“the proportion of cows with an intramammary infection (IMI), where infection was defined as LS ≥4.0; the proportion of cows with a new IMI, where new IMI was defined as LS changing from <4.0 to ≥4.0 in the last 2 tests; the proportion of cows with a chronic infection, where chronic was defined as LS ≥4.0 on the last 2 tests…”*

Dufour, S., Dohoo, I.R., Barkema, H.W., Descôteaux, L., Devries, T.J., Reyher, K.K., Roy, J.-P., Scholl, D.T., 2012. Manageable risk factors associated with the lactational incidence, elimination, and prevalence of Staphylococcus aureus intramammary infections in dairy cows. Journal of Dairy Science 95, 1283–1300.

In the above study, linear regression was used to model herd prevalence: *“The relative effect of herd incidence and elimination rates on the herd prevalence of IMI was assessed. For this purpose, a linear regression model was used. The dependent variable was the quarter prevalence of Staph. aureus IMI in the herd over the 2 yr of the study, and explanatory variables were the herd incidence and elimination rates during that same period.”*

Lobeck, K.M., Endres, M.I., Shane, E.M., Godden, S.M., Fetrow, J., 2011. Animal welfare in cross-ventilated, compost-bedded pack, and naturally ventilated dairy barns in the upper Midwest. Journal of Dairy Science 94, 5469–5479.

In the above work, *“A linear mixed model (MIXED procedure, SAS Institute Inc.) was built to evaluate the association between housing system and the outcome variables: lameness prevalence, hock lesion prevalence, BCS, hygiene score, CCI, SUI, mortality rate, turnover rate, respiration rates, and mastitis infection prevalence,”* where “mastitis infection prevalence” was defined as “c*alculated by the number of animals with a test SCC >200,000 cells/mL divided by the total number of animals in the pen.”*

Eckelkamp, E. A., J. L. Taraba, K. A. Akers, R. J. Harmon, and J. M. Bewley. 2016a. Sand bedded freestall and compost bedded pack effects on cow hygiene, locomotion, and mastitis indicators. Livestock Science 190:48-57.

The above work used “*The MIXED procedure of SAS … to develop all models for all barn type comparison analyses…High SCC prevalence and RCMI explanatory variables were barn type, maximum ambient THI period, mean herd hygiene score, and all 2-way interactions. High SCC prevalence and RCMI were calculated as a herd percentage for each visit period. High SCC prevalence was the % of the herd with a DHI SCC ≥200,000cells/mL at each visit period.”*

***Dpoe sthis rule pout a biologically important negative effect [sic] (referencing %newly elevated SCC across facility groups)***

***Again does this rule out a biolgically important negative effect [sic] (referring to % chronically elevated SCS coefficient estimates by facility type)***

***Again prevalence data (referring to % SCS ≥ 4.0 current test)***

***Again doesn't rule out biologically important negative effect (or positive effect)***

AU: See previous comment.

***Reviewer: 2***

***Major (General or Overall) Comments to the Author. Should be concrete and helpful for revision and may include overall strengths and weaknesses of the scientific merit and research approach. Authors ARE expected to reply to these comments.***

***Dear authors, the paper is interesting but I think a revision is needed. See my suggestions below:***

AU: Thank you for your input and thoughtful reading of our manuscript. Responses to all suggestions are below.

***Line 6: Similar for what?***

AU: Thank you for requesting clarification on this point. Language was added to specify that bedded packs had similar milk quality, udder health, udder hygiene and milk production to tiestalls and freestalls in the current study. We also have clarified that all three systems are achieving acceptable levels of milk quality in the context of bedding on organic materials.

***Line 53: All farms or only those with bedded packs? That’s confusing***

AU: Thank you requesting clarification on this point. Specification was added (now line 61) to clarify that these finding were from thesecondary analyses which combined all 21 farms from the three housing categories together.

***Line 90: previous works?***

AU: Thank you for this question; we interpret this as asking us to clarify what findings from the literature we are referring to in this sentence. The basis for the conjecture that mastitis risk could possibly be higher for cows housed on bedded packs is detailed in the following sentences after line 90, with references to literature supporting each component of the conjecture (lines 92-105): (1) loose-housed cows continually add manure to the bedded pack, which contributes both pathogenic bacteria and nutrients to the organic bedding material; and (2) organic bedding material is more likely to have a higher bacteria count than inorganic bedding, such as sand. Organic bedding has been shown to be inherently associated with a higher number of bacteria on teat ends, and higher concentration of bacteria in general in bedding is associated with a higher concentration of bacteria on teat ends. We therefore conclude that theoretically, this higher concentration of bacteria on teat ends may put the mammary gland at an increased risk of infection, although clearly acknowledge that evidence to support this relationship is weak or lacking (lines 102-105).

***Line 95: Bacterial***

AU: Thank you for this suggestion, and this edit has been made at line 96.

***Line 97-101: I don’t appreciate the difference between reason 1 and 2.***

AU: The first two reference work which compared bacterial counts on teat ends for cows bedded with an inorganic material vs. an organic material, and found the teat ends of cows on organic bedding had a higher bacterial count than those bedded on an inorganic substrate*.* The second group of references are work that established that a higher concentration of bacteria in bedding (generally) is related to a higher concentration of bacteria on teat ends. As bedded packs are both organic bedding and can have a high bacterial count in the material (lines 92-98), these two separate points were made to support the assertion that cows on bedded packs have the potential to have higher concentrations of bacteria on teat skin. We have added a few words to these lines that hopefully clarify this nuance.

***Line 111: This “actively-managed composting bedded packs” is something very (too?) specific. Do you mean that the bedded pack barns included have some particular management procedures aimed at promoting the composting process? Or are these just “normal” (US-type) compost-bedded pack barns? In any case, further explanation of what CBP are and the different management styles (composting vs non-composting if you want to keep using that specific term) is lacking in the introduction. This is particularly important as the whole paper focuses on housing systems and CBP in particular.***

AU: Thank you for requesting clarification around the terminology of “bedded pack” throughout the manuscript. We appreciate that the term “bedded pack” is vague, as it encompasses multiple styles of management and appears to have variable use in the literature and lay-publications. For the purpose of this study, we use the general term “bedded pack” to include both “conventional bedded packs” and “compost bedded-packs,” as defined in Bewley et. al (2017). To clarify this for the reader, we have defined what we mean by the term “bedded pack” in the Introduction from line 82-97, and further delineate this grouping when describing the enrolled herds (Materials and Methods, lines 185-208). This section also attempts to more clearly delineate composting bedded packs and conventional bedded packs, which hopefully addresses your suggestion to further explain the different types of management for bedded packs. We acknowledge that this was missing from the Introduction, and is an important revision. We hope this clarifies our use of the term, and the diversity of management styles under this umbrella currently being used by dairy producers in the Northeast US.

***Line 116-118: I’m not sure about this statement. Are you? And is this important for the paper?***

AU: Thank you for this question. More language was added to the manuscript to clarify what was meant by this statement. Although a formal, systematic literature review was not performed, the review of literature available from peer-reviewed publications in the English language around milk quality, udder hygiene, and mastitis risk and bedded pack housing for lactating dairy animals was quite extensive. The three previously mentioned publications (line 111-116) compared compost bedded packs to freestall facilities, and as the manuscript currently reads (“to the best of our knowledge”), no peer-reviewed publications in the English language were encountered that directly compared these specific metrics (udder hygiene, aerobic milk culture data, and mastitis risk) between tiestall herds and herds using a bedded pack system. As this is an aspect of the current study that is a novel contribution to bedded pack research, we thought it was important to highlight it as such. There are two key industry issues that we think make this statement important to include in the manuscript. 1) Globally, the dairy industry is under increasing pressure to phase out the use of tiestall confinement facilities, forcing individual dairy farmers to seek alternative housing systems. 2) In seeking alternative housing systems, dairy farmers are likely to be interested in bedded pack systems as an alternative to freestall barns for the reasons cited in our manuscript and the literature. These reasons suggest to us that reporting milk quality metrics on tiestall herds with those adopting bedded pack systems in the same geographic region is valuable data.

***Line 120: Bedded pack barn was already abbreviated.***

AU: Thank you for this suggestion. The term “compost bedded-pack” was abbreviated as “CBP” at line 111 in the Introduction, specifically to refer to this type of system in previous work on systems which are actively-managed for aerobic composting. This was done to distinguish actively-managed, compost bedded-packs from our use of the more general term “bedded pack (BP)” throughout the rest of the manuscript; which, at this point in the manuscript, we had not yet defined and abbreviated.

We have now added a definition of how we use the term “bedded pack” earlier in the manuscript (see previous reply to comment), and define our abbreviation of the term as “BP” at line 98. Abbreviations (also for tiestall and freestall) have now been updated and edited throughout the manuscript, as suggested below.

***Line 123: Bedded pack barn was already abbreviated.***

AU: See previous comment.

***Line 128: Bedded pack barn was already abbreviated.***

AU: See previous comment.

***Line 156: Still a bit of confusion with “bedded pack” and its abbreviation. Please revise the use of abbreviations in the whole manuscript.***

AU: See previous comment.

***Line 160: “convenience sample”?***

AU: Thank you for your question about our use of this term, and we interpret that this comment is a request to extrapolate on what we meant by using it. A convenience sample is a commonly-used term describing methodology of epidemiological studies, defined as “a sample of study subjects selected for expedience … usually chosen because they are readily available” (Oleckno, W. A. 2008. *Epidemiology: Concepts and Methods*. Waveland Press). By using this term here, we are being transparent that participating farms were not enrolled using any method of random selection; with the relatively small number of organic dairy farms in Vermont using a bedded pack system (as indicated by the survey administered, described in a previous publication, Andrews et al., 2021), random selection to enroll farms would not have resulted in a sufficient number of farms using this housing style of interest for lactating dairy cows. We acknowledged that convenience sampling can result in selection bias (lines 724-726), and describe attributes of the study population in regards to the source population in order to inform the reader how representative the enrolled farms may be of the larger source population (lines 727-730).

***Line 179-191: So each farm enrolled was visited once during the period? Please explain.***

AU: Thank you requesting clarification on this point, as it is not explicitly stated in the manuscript. Each herd was visited once during the study period, and specification of this was added on line 183.

***Line 191: Again, please revise the use of abbreviations throughout the paper.***

AU: See previous comment.

***Line 214: You should specify how the bedding samples were collected and analysed. At what depth? Single sample taken at one single location or composite sample taken from multiple points across the pen? How the bedding samples were transported/treated/analysed…? Hopefully bedding samples have been analysed for bacterial concentration(s) or something else (otherwise what was the reason to get them?). This is completely unclear… Also other pack measures (like temperature) deserve more detailed explanation. Consider adding a section dedicated to pack measures.***

AU: Thank you requesting clarification on this point. Bacterial analysis of bedding samples and pack temperature measurements were completed for these farms, but are outside the scope of the current manuscript. As such, we acknowledge that it is appropriate to remove mention of them from the current manuscript (lines 215, 225-228, 232).

***Line 228-229: Not clear how bedding depth was measured. Please revise.***

AU: Thank you requesting clarification on this point. Additional information as to how the bedded pack depth was measured was added to lines 269-271.

***Line 297: CFU counts of what? Bulk tank milk? Please specify***

AU: Thank you requesting clarification on this point. Specification that this was referring to cfu counts of bacteria from bulk tank milk was added to line 299.

***Line 299 and 318 and 321: “Offering” variables to a model sounds awkward (at least to me).***

AU: Thank you for this observation. “Offering” variables to a model is a widely-used phrase, and we are opting currently to keep this language in the manuscript. Briefly, a few examples from publications that use this term in the *Journal of Dairy* Science include:

Bicalho, R. C., V. S. Machado, and L. S. Caixeta. 2009. Lameness in dairy cattle: A debilitating disease or a disease of debilitated cattle? A cross-sectional study of lameness prevalence and thickness of the digital cushion. J Dairy Sci 92(7):3175-3184.

Patel, K., S. M. Godden, E. Royster, B. A. Crooker, J. Timmerman, and L. Fox. 2019. Relationships among bedding materials, bedding bacteria counts, udder hygiene, milk quality, and udder health in US dairy herds. J. Dairy Sci. 102(11):10213-10234.

Pires, J. A. A., T. Larsen, and C. Leroux. 2022. Milk metabolites and fatty acids as noninvasive biomarkers of metabolic status and energy balance in early-lactation cows. J. Dairy Sci. 105(1):201-220.

Rowe, S. M., S. M. Godden, E. Royster, J. Timmerman, B. A. Crooker, and M. Boyle. 2019. Cross-sectional study of the relationships among bedding materials, bedding bacteria counts, and intramammary infection in late-lactation dairy cows. J Dairy Sci 102(12):11384-11400.

***Line 334-340: Not completely clear how the analysis for Objective 2 was performed. Just linear regression between each pair of variables? Straight “stat” package? What’s that P-value at line 339?***

AU: Thank you for this request for clarification. The linear regression for Objective 2 was performed exactly as previously described for Objective 1 in the Materials and Methods section (lines 351-355), but this was not explicitly stated before and has been added at line 377. Univariate linear regression was performed in R using the “lme4” package to investigate the unconditional relationship between the eight outcomes of interest and the various herd-level independent variables. Language clarifying that the significance level of P ≤0.20 was for an F-test was also added.

***Line 426: Please explain briefly how facility type affected avg SCS.***

AU: text was added describing this at lines

***Line 466-474: Was this already discussed? Or was this section specifically intended for bedded-pack barns? That’s unclear and sounds like a repetition. Please revise this section. Further, and perhaps more importantly, bedding depth seems the most important predictor of udder hygiene and possibly mastitis risk but how this variable is defined remains very unclear (just some sparse info in the supplementary material). I strongly suggest adding more info about how bedding depth was measured (in the different housing systems) and describe its distribution in the results (mean, SD and range?).***

AU: This section describes findings for three separate groupings of facility type (not a repetition), each of which is specified in the description of the results (“*herds using a bedded pack,”* originally line 467; “*cows housed on some type of deep bedding (i.e., grouping all herds reporting deeply-bedded stalls plus bedded pack herds,”* originally line 469; “*for the* *fifteen farms reporting bedding depth in stalls,”* originally line 472). We have taken this suggestion and added language that hopefully clarifies that these are three separate groupings.

Description of how bedding depth was measured is included on lines 269-273 in the Materials and Methods, and was elaborated as per the previous suggestion. We agree that adding this information to the Results section improves the manuscript; this has been done on lines 409-412.

***DISCUSSION SECTION: To me, this section feels way too long, not always relevant and with some repetitions. I’d consider shortening it sensibly and verify the concepts discussed as well as other studies mentioned are closely related to the results of the current study.***

AU: Thank you for this suggestion; we have edited the original content of the Discussion section to make it shorter.

***Line 490-492: I don’t feel the need to add this conclusive statement here. The reader might lose interest in the upcoming discussion section.***

AU: Thank you for this suggestion; we have removed this sentence from the brief summary preceding the Discussion section which JDS requests.

***Line 498-501: This was already mentioned several times. No need to highlight it again.***

AU: Thank you for this suggestion; this sentence was removed.

***Line 626: Revise tense structure***

AU: Thank you for this suggestion; this edit was made.

***Line 677-682: Revise use of brackets***

AU: Thank you for this suggestion; this edit was made.

***Line 706-721: To me, merging “static” bedded-pack barns (could also be defined as “conventional straw yards”?) and compost barns in the same group has several limitations as these systems are known to produce different results, especially regarding udder health. Although this issue is somehow justified by the relatively small number of farms available/included I think it deserve further discussion.***

AU: Thank you for this comment, and we agree that this point should be more overtly addressed in the manuscript. Other reviewers requested clarification on this as well; please see response to comment for line 350 from Reviewer 1.

***Reviewer: 3***

***Major (General or Overall) Comments to the Author. Should be concrete and helpful for revision and may include overall strengths and weaknesses of the scientific merit and research approach. Authors ARE expected to reply to these comments.***

***This is an interest study though limited in its description of the sample population and the number of herds in each category of farms. In some ways, it's almost a case study.***

AU: Thank you, we agree it is almost a case study, but does not exactly fit that definition as we understand it. We believe it is more accurately described as a non-probability, observational (descriptive), analytical cross-sectional study where the unit of interest is the herd and we examine “the association between exposures and outcomes on a snap of time.” (Rezigalla, 2020). Alternatively, Crowe et al. (2011) provide a review broadly defining a case study as “an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context” that includes “both the process of learning about the case and the product of our learning.” They identify a “collective case study” as a particular category of case studies, which involves studying multiple cases simultaneously or sequentially in an attempt to generate a broad appreciation of an issue or phenomenon. These two definitions of case studies are not incompatible. We believe our observational (or descriptive) study rises above the definition of a collective case study, as we set out to assess associations guided by a hypothesis and our results can also be seen as hypothesis-generating (Rezigalla, 2020); i.e., we had a prior hypothesis, which was to compare different management practices in a defined sample of herds.

***Line 108: Is this referring to non-composted bedded pack barns? Perhaps to minimize confusion that should be clarified with a term like "traditional bedded pack barns" or "non-aerated compost bedded pack barns."?***

AU: Thank you for this important suggestion requesting clarification around the terminology of “bedded pack.” Other reviewers also asked for clarification (see previous comments for more detailed response). Briefly, we acknowledge that a clearer description of different kinds of bedded packs and how we used the term was missing from the Introduction, and is a significant revision. We appreciate that the term “bedded pack” is somewhat vague, as it encompasses multiple styles of management and appears to have variable use in the literature and lay-publications. For the purpose of this study, we use the general term “bedded pack” to include both “conventional bedded packs” and “compost bedded-packs,” as defined in Bewley et. al (2017). To clarify this for the reader, we have now defined what we mean by the term “bedded pack” in the Introduction from line 82-97, and further delineate this grouping when describing the enrolled herds (Materials and Methods, lines 185-208). This section also attempts to more clearly delineate composting bedded packs and conventional bedded packs. We hope this clarifies our use of the term, and the diversity of management styles under this umbrella currently being used by dairy producers in the Northeast.

***Line 120: What months comprise the non grazing season?***

AU: Thank you requesting clarification on this point. Specification of the term “non-grazing season” for Vermont was added to lines 5 (Interpretive Summary), 22 (Abstract), and 121 (Introduction).

***Line 230. How were the cows randomly selected?***

AU: Thank you requesting clarification on this point. Random selection was not actually used to select these cows; the first 30 cows that were able to be evaluated in a loose pen were included, or the first 30 encountered in a tiestall. Language correcting this point has been added (lines 269-270).

***Lines 350-352. The mix of cultivated and non-cultivated bedded pack barns is a bit concerning. This is the first mention of this. This should be indicated within the abstract and earlier discussion. This is similar to having both sand and organic bedding freestall barns in the same category. Both scenarios are grouping herds that ideally would be different sample groups as they are considerably different.***

AU: Thank you for this comment, and we agree that this point should be more overtly addressed in the manuscript. Other reviewers requested clarification on this as well; please see response to comment for line 350 from Reviewer 1.

***Line 493: It may be worth discussing the production level of these herds was low which might have reduced mastitis susceptibility.***

AU: Thank you bringing up this interesting point. These farms had comparatively higher-producing cows when compared to other organic farms in Vermont (lines 774-776), but certainly when compared to the general population of dairy cows in the U.S. they were low-producing. Language acknowledging this point has been added to the manuscript (lines 776-780).

***The sample size is a limiting factor in this study. It would be worthwhile to add some discussion of some of the potentially biologically significant differences observed in the data a bit more (i.e. tie-stall barn cows hygiene scores were considerably higher.)***

AU: Thank you for this comment, which other reviewers also noted. See our responses to reviewer 1 above.

Astiz, S., F. Sebastian, O. Fargas, M. Fernández, and E. Calvet. 2014. Enhanced udder health and milk yield of dairy cattle on compost bedding systems during the dry period: A comparative study. Livestock Science 159:161-164.

Benson, A. F. 2012. Consider deep pack barns for cow comfort and manure management. Accessed March 18, 2024. Cornell University, Ithaca, NY. <https://smallfarms.cornell.edu/2012/04/consider-deep-pack-barns-for-cow-comfort-and-manure-management/>.

Bewley, J. M., L. M. Robertson, and E. A. Eckelkamp. 2017. A 100-Year Review: Lactating dairy cattle housing management. J. Dairy Sci. 100(12):10418-10431.

Fregonesi, J. A. and J. D. Leaver. 2001. Behaviour, performance and health indicators of welfare for dairy cows housed in strawyard or cubicle systems. Livestock Production Science 68(2):205-216.

Fregonesi, J. A. and J. D. Leaver. 2002. Influence of space allowance and milk yield level on behaviour, performance and health of dairy cows housed in strawyard and cubicle systems. Livestock Production Science 78(3):245-257.

Leso, L., M. Barbari, M. A. Lopes, F. A. Damasceno, P. Galama, J. L. Taraba, and A. Kuipers. 2020. Invited review: Compost-bedded pack barns for dairy cows. J Dairy Sci 103(2):1072-1099.

Oleckno, W. A. 2008. Epidemiology: Concepts and Methods. Waveland Press.

Peeler, E. J., M. J. Green, J. L. Fitzpatrick, K. L. Morgan, and L. E. Green. 2000. Risk Factors Associated with Clinical Mastitis in Low Somatic Cell Count British Dairy Herds. J. Dairy Sci. 83(11):2464-2472.

Shepley, E., J. Lensink, H. Leruste, and E. Vasseur. 2020. The effect of free-stall versus strawyard housing and access to pasture on dairy cow locomotor activity and time budget. Applied Animal Behaviour Science 224:104928.

Ward, W. R., J. W. Hughes, W. B. Faull, P. J. Cripps, J. P. Sutherland, and J. E. Sutherst. 2002. Observational study of temperature, moisture, pH and bacteria in straw bedding, and faecal consistency, cleanliness and mastitis in cows in four dairy herds. Vet Rec 151(7):199-206.

Andrews, T., C. E. Jeffrey, R. E. Gilker, D. A. Neher, and J. W. Barlow. 2021. Design and implementation of a survey quantifying winter housing and bedding types used on Vermont organic dairy farms. J. Dairy Sci. 104(7):8326-8337.

Whistance, Lindsay et al. “Defaecation behaviour of dairy cows housed in straw yards or cubicle systems.” Applied Animal Behaviour Science 105 (2007): 14-25. DOI: 10.1016/j.applanim.2006.05.010

Rezigalla AA. Observational Study Designs: Synopsis for Selecting an Appropriate Study Design. Cureus. 2020 Jan 17;12(1):e6692. doi: 10.7759/cureus.6692. PMID: 31988824; PMCID: PMC6970097.