J. Dairy Sci. 107:11774–11784
https://doi.org/10.3168/jds.2024-25163

Enhancing evaluation of bull fertility through multivariate analysis of sperm

H. C. Azevedo,[1] H. D. Blackburn,[2]* E. A. Lozada-Soto,[2] S. F. Spiller,[2] and P. H. Purdy[2]
[1]Brazilian Agricultural Research Corporation (Embrapa), Aracaju 49025-040, SE, Brazil
[2]U.S. Department of Agriculture–Agricultural Research Service (USDA ARS), National Animal Germplasm Program, Fort Collins, CO 80521

ABSTRACT

Computer-assisted sperm analysis (CASA) has become the predominant tool for assessing bull semen in AI programs. Despite such popularity CASA's ability to predict fertility has been limited, especially when emphasis is based on single motion characteristics. Our hypothesis is that numerical sets of CASA measures provide a more effective method to differentiate the potential fertilization capacity of bulls and that bulls can be clustered based on sets of CASA measures. Therefore, we used CASA to evaluate frozen-thawed semen samples from 307 Holstein and 152 Jersey bulls sourced from USDA Agricultural Research Service's National Animal Germplasm Program gene bank. Sperm was evaluated immediately after thawing and 30 min later. We evaluated sperm kinetic and morphometric means and variances to capture the structure of CASA data in relation to various sources of variation. These data were subjected to univariate and multivariate statistical methods to investigate animal and management factors affecting sperm characteristics measured by CASA. Clustering with K-means identified 4 clusters of bulls based upon each cluster's set of CASA parameters after thawing. There was little overlap among clusters for sets of CASA measures. At the extremes, bull cluster 1 (BC1, n = 180) and BC3 (n = 101) had different sire conception rates (SCR) −0.07 versus −1.29, respectively, and sets of CASA measures. Interestingly, bull cluster 2 (BC2) had CASA measures that could be perceived as negative, for example, cell size at 8.18 mm^2 versus 6.37 mm^2 for bull cluster 4 (BC4) and total motility of 29.7% versus 48.7% for BC3, but SCR for BC2 were higher (−0.79) than those for BC3 (−1.29). Despite such discrepancies for some BC2 CASA values it appears the potentially negative effects were offset by the levels of other CASA values. Our findings suggest improved approaches for using CASA could lie in evaluating multiple CASA measures as sets within specific numerical ranges rather than as independent measures.

Key words: cattle, cryopreserved semen, post-thaw evaluation

INTRODUCTION

Currently, there is no definitive method to assess post-thaw sperm fertility among mammalian species like cattle. In part, this is due to a lack of clarity in sperm variability but also insemination protocols, cow physical and reproductive condition, and the likelihood of successful fertilization. However, livestock gene bank managers need a better way of assessing the fertilizing potential of samples stored in repositories to adjust collection sizes and their potential use. In general, the low correlation between measures like motility and progressive motility and fertility, although widely used, lacks explanatory power, especially as methodologies for sperm viability and genetic structure have advanced.

The growing importance of computer-assisted sperm analysis (**CASA**) is evident in its role in quality control for semen used in cattle AI. However, it is not obvious that information generated by CASA is being used to its full potential and an overdependence on the motility and progressive motility parameters has been suggested, especially given their low capability to predict fertility (Amann and Waberski, 2014). The affordability of CASA is increasing and there is a need to move beyond using it as a tool for measuring motility and progressive motility and toward a more comprehensive package of analyses that includes flow cytometric and morphological analyses. Such a shift requires more robust assessments to help foster the development of new tools (e.g., free software; Alquézar-Baeta et al., 2019), artificial intelligence enhancements (Ehlers et al., 2011), and smartphone applications (Park et al., 2021).

Computer-assisted sperm analysis has been used to identify some movement patterns of bull sperm with high versus low fertilization rates (Verstegen et al., 2002). However, concerns have been raised about the ambiguity of CASA's parameters to predict fertility (Silva et al., 2023) and their apparent lack of correlation when assessed separately (Kasimanickam et al., 2006). Therefore, debates exist regarding the correlation be-

Received May 13, 2024.
Accepted August 16, 2024.
*Corresponding author: harvey.blackburn@usda.gov

The list of standard abbreviations for JDS is available at adsa.org/jds-abbreviations-24. Nonstandard abbreviations are available in the Notes.

tween CASA-measured sperm kinetic and morphological parameters and in vivo fertility rates (Farrell et al., 1998; Li et al., 2016; Zoca et al., 2023).

In addition, many entities only scrutinize the average CASA values from analysis, resulting in a very limited interpretation of the results because the means mask variability among and within individuals, a key for assessing fertility (Amann and Waberski, 2014). According to Jeong and Chong (2020), although mean and variance values summarize characteristics, their combined use is essential for understanding the magnitude of a dataset. Therefore, using the mean and variance from CASA analyses will become more valuable when we understand the relationship of CASA parameters from a bull or a sample.

In this paper, we use univariate and multivariate statistical approaches to evaluate combining various CASA parameters as predictors of fertility. The cluster analysis used can reveal populations of sperm with similar features and potentially improve the accuracy of fertility assessments by delving deeper into the complex data derived from the analyses (Rodriguez et al., 2019). In addition, we depart from relatively small numbers of matings by using the dairy industry's standard measure of bull fertility, sire conception rate (**SCR**), which is computed across hundreds if not thousands of matings per bull and adjusts for other factors related to the cow, bull, herd, management, season, nutrition, animal age, reproductive success, and environmental factors (CDCB, 2024), thus making this breeding value quite robust. Therefore, this study aimed to evaluate dairy bull sperm kinematic and morphometric CASA values, characterize, and stratify bulls into clusters, and then associate bull clusters with measures of sperm quality and bull fertility (SCR). The overarching goal was to develop a more robust model, using CASA parameters and SCR values, for assessing fertilizing potential.

MATERIALS AND METHODS

This work did not require ethical approval because it did not use animals but only semen doses from commercial AI companies (studs) for its execution.

Animal and Semen Sample Information

We analyzed cryopreserved semen from 307 Holstein (**HO**) and 152 Jersey (**JE**) bulls, sourced from the USDA Agricultural Research Service's National Animal Germplasm Program. The birth years of the bulls ranged from 1994 to 2016. All these bulls were originally sourced from commercial AI companies (studs) and it was their decision as to how many bulls to contribute. The bulls were managed under the collection strategies of each AI stud, resulting in each bull's maximal sperm quality and production for marketing purposes.

Data obtained about each bull was combined into a common file and included: bull ID, breed, associated AI center (AI stud), and birth year, sourced from the Animal Germplasm Resources Information Network of the National Animal Germplasm Program. The SCR for each bull was added, as were post-thaw CASA parameters.

Sire Conception Rates

The SCR of bulls was obtained from the Council on Dairy Cattle Breeding databases accessed between May and July 2023 (CDCB, 2024). The SCR is an evaluation of a bull's fertility under normal management conditions. It is calculated using AI data and considers the probability of a cow becoming pregnant from a single insemination. The SCR is calculated as the difference between the observed success rate and the expected success rate, adjusted for various sources of variation, such as cow age, reproductive history, herd conditions, reproductive management, cow health and nutrition, environmental conditions, and season of the year (CDCB, 2024). It also considers factors including inbreeding, the bull's age, the AI organization, mating year, and the bull's effect (Norman et al., 2008). By using each bull's SCR as the measure of fertility, these issues are accounted for, thereby reducing bias in the analysis.

Sperm Analysis (CASA)

Frozen semen samples were thawed at 37°C for 30 s and diluted in Tris-buffered medium (200 m*M* Tris, 65 m*M* citric acid, 55 m*M* glucose; Purdy and Graham, 2004). Because it has been suggested that sperm damage may not be fully visible immediately after thawing (Castro et al., 2016), we evaluated semen under incubation at 37°C at 2 post-thaw time points: **T1** immediately after thawing; and **T2**, 30 min after the first evaluation. The T1 time point corresponded to 7 min of incubation, which is within the recommended range (5–10 min.) for sperm motility to begin to be fully expressed (Barth, 1989).

Semen samples were analyzed using a Hamilton Thorne Motility Analyzer (Version 14 IVOS, Beverly, MA) referred to in this work as CASA. The CASA was configured with the following settings: 30 frames per second were captured at a frame, capture rate of 60 Hz; minimum contrast: 80; minimum cell size: 5 pixels; cutoff for average path velocity (**VAP**): 30 μm/s; minimum cutoff for progressive motility: 50 μm/s of VAP; cutoff for straight-line velocity (**VSL**): 15 μm/s; threshold straightness (**STR**; VSL/VAP): 70%; static head size ranging from 0.53 to 4.45; enlargement factor: 1.89. A

minimum of 700 cells was analyzed in at least 5 fields (Purdy and Graham, 2004).

Spermatozoa were assessed for kinematic parameters: VAP (μm/s), VSL (μm/s), curvilinear velocity (**VCL**, μm/s), amplitude of lateral head displacement (**ALH**; μm), beat-cross frequency (**BCF**; Hz), STR (%), and linearity (**LIN**; VSL/VCL, %), and 2 morphometric parameters, elongation ratio (**ELO**; ratio of the sperm head width to the sperm head length, %) and head size (**HSZ**, μm^2). The mean and variance of these parameters for each bull, as well as the overall sample motility parameters (motile sperm, **MOT**, %; progressive motile sperm, **PMOT**, %), were recorded.

Statistical Analysis

Descriptive Statistics. Data of descriptive statistics of sperm kinematic and morphometric, and sample motility parameters at T1 and T2 are in Supplemental Tables S1 and S2 (see Notes).

Correlation Analysis. The complete dataset was used for Pearson and Spearman correlation analyses at T1 and T2.

Univariate Analysis. We used a reduced dataset for this analysis in which we removed bulls from AI studs with less than 5 animals (n = 4) or bulls with samples without a known evaluation date (n = 20). Thus, the bulls were distributed among the studs as described in Table 1.

To estimate the effect of genetic, management, and sperm analysis factors on sperm kinematic and morphology and sample motility parameters we fit the following linear mixed model using the lme4 R package (Bates et al., 2015):

$$y_{ijklmn} = \mu + Breed_i + Stud_j + Breed \times Stud_{i:j} + ET_k + BY_l + ED_m + ID_{n(i:j)} + e_{ijklmn},$$

where y_{ijklmn} is the response variable corresponding to a sperm parameter; μ is the overall mean (intercept); $Breed_i$ is the fixed effect of the ith level of breed (i = HO, JE); $Stud_j$ is the fixed effect of the jth AI stud (j = StudA, StudB, StudC, StudD); Breed × $Stud_{i:j}$ is the fixed interaction term of breed i in stud j; ET_k is the fixed effect of the kth level of evaluation time point (k = T1, T2); BY_l is the linear covariate of year of birth; ED_m is the random effect of the mth sample evaluation day; $ID_{n(i:j)}$ is the random effect of the nth bull nested within breed i and stud j; and e_{ijklmn} is the random residual effect. Random effects were assumed to have normal distributions with means of 0 and variances equal to the estimated variances σ^2_{ED}, σ^2_{ID}, and σ^2_e for ED, ID, and e, respectively. An ANOVA was performed, and *F*-values and *P*-values were obtained for all effects. Additionally, to estimate the average value of sperm parameters at each level of fixed effects we obtained least squares means estimates using the emmeans R package (v.1.8.8; Lenth, 2023).

Multivariate Analysis. We hypothesized that bull fertility (as measured by SCR) is dependent on more than one CASA parameter being in a viable range; furthermore, there is an interdependence of parameters upon one another. Due to the interdependence of observed variables, it should be possible to develop sets of variables with greater explanatory power for which principal component (**PC**) and exploratory factor analyses are well suited. Therefore, PC analysis (**PCA**) was performed on the complete dataset using the R "prcomp" function based on the sperm kinematic parameters at the 2 time points. Afterward, to reduce the number of PC to those of the biggest explanatory power for downstream analyses, we used the Kaiser criterion to only keep those with an eigenvalue greater than or equal to 1.00. The matrix of eigenvectors and the proportion of variance explained by each PC were obtained.

To identify groups of bulls with similar sperm characteristics we obtained clusters based on the reduced set of PC. We performed a nonhierarchical clustering procedure using K-means on the Euclidean distance matrix built from the scores of the reduced set of PC. The optimal number of bull clusters (**BC**) was determined by visualization of the average silhouette width, the gap statistic, and the total sums of squares within each cluster (k), where k ranged from 1 to 10 clusters. Based on these statistics we selected 4 and 2 to be the appropriate number of BC for T1 and T2, respectively. Potential differences in sperm kinematic and morphometric parameters, sample motility parameters, and SCR between clusters were assessed using univariate models with cluster membership as the sole explanatory variable. Least squares means estimates and 95% CI were obtained for each bull cluster.

All statistical analyses were done using the R statistical software (v.4.3.1; R Core Team, 2023). Statistical significance was considered at α = 0.05.

Table 1. Number of bulls by breed from each stud

	Stud			
Breed	A	B	C	D
Holstein	254	11	19	20
Jersey	124	3	9	4
Total	378	14	28	24

RESULTS

Factors Influencing CASA Parameters

Table 2 shows the ANOVA *F*-values for the effects of breed, AI stud, breed × AI stud, and evaluation time on

Table 2. ANOVA *F*-values for the main effects of breed, AI stud, the interaction effect between breed and AI stud, and sample evaluation time point on sperm kinematic and morphometric parameters

Sperm parameter[1]	Breed	AI stud	Breed × AI stud	Evaluation time point
Kinematic (mean)				
VAP-m	0.65	43.51*	1.58	4.23*
VSL-m	0.00	53.03*	0.55	4.63*
VCL-m	2.29	46.16*	2.14	2.54
ALH-m	6.47*	34.52*	2.93*	0.78
BCF-m	6.70*	8.84*	0.66	26.67*
STR-m	2.51	22.35*	2.23	2.54
LIN-m	5.94*	17.94*	1.85	5.72*
Kinematic (variance)				
VAP-v	0.28	16.31*	6.75*	6.33*
VSL-v	0.05	13.61*	3.26*	14.24*
VCL-v	1.35	25.42*	7.79*	21.28*
ALH-v	2.20	22.42*	6.50*	71.13*
BCF-v	7.16*	23.54*	6.63*	18.92*
STR-v	1.22	24.43*	1.44	105.90*
LIN-v	5.32*	46.24*	7.14*	162.24*
Morphometric (mean)				
ELO-m	36.72*	6.48*	8.14*	9.83*
HSZ-m	7.01*	25.05*	4.25*	0.22
Morphometric (variance)				
ELO-v	24.26*	4.88*	14.40*	1.82
HSZ-v	1.40	23.75*	3.85*	0.07

[1]Sperm parameter: VAP = average path velocity, VSL = straight-line velocity, VCL = curvilinear velocity, ALH = amplitude of lateral head displacement, BCF = beat-cross frequency, STR = threshold straightness, LIN = linearity, ELO = elongation ratio, HSZ = head size.

*Denotes an estimate with a *P*-value smaller than 0.05.

sperm parameters. Figure 1 and Figure 2 show the LSM estimates of sperm parameters by breed and AI stud and semen evaluation time, respectively.

Breed significantly affected the mean (**m**) of ALH-m, BCF-m, and LIN-m, and the variance (**v**) of BCF-v and LIN-v. The effect of the AI stud was significant ($P < 0.05$) for all kinematic parameters, and we speculate that this may be caused by such things as differences in cryopreservation media components, initial sperm quality, or technician variability. The interaction between breed and AI stud was significant for ALH-m and all the variances of kinematic parameters except for STR-v. We found a significant effect of breed for all morphometric measures except for HSZ-v. Both the AI stud and the breed × AI stud significantly affected the mean and variance of all morphometric sperm parameters measured.

The effect of breed on sperm parameters was most prominent for bulls in Stud A. The JE bulls had significantly larger VCL-m, ALH-m, VAP-v, VCL-v, ALH-v, and STR-v and had significantly lower STR-m and LIN-m estimates than HO bulls. We also observed significant differences for several kinematic and morphometric parameters between bulls in Stud D. HO bulls had significantly larger ELO-m, HSZ-m, BCF-v, LIN-v, and ELO-v and nonsignificant LIN-m than JE bulls. Bulls within Stud A had higher estimates of the means of velocity parameters than other studs. For STR-m and LIN-m, we observed significantly lower estimates in Stud D compared with the remaining studs. On the other hand, Stud D had significantly larger sperm (HSZ-m) than the other studs. Across all measures of sperm parameter variance, we observed the largest estimates for Stud D, significantly higher for ALH-v, BCF-v, STR-v, LIN-v, ELO-v, and HSZ-v. Figure 1 suggests that JE and HO samples change ranking especially when comparing studs B and D and therefore this reranking may have caused the interaction.

The semen post-thaw evaluation time point had a significant effect on VAP-m, VSL-m, BCF-m, LIN-m and all measures of variance of kinematic parameters, and ELO-m (Table 2). We found a significantly lower variance estimate for STR-v and LIN-v when comparing T2 and T1 (Figure 2).

Relationship Between CASA Parameters and SCR

The correlation analysis results between kinematic and morphometric sperm parameters and SCR at T1 and T2 are presented as heat maps in Supplemental Figures S1 and S2 (see Notes). In all instances, the correlations between CASA readings and SCR were less than +0.20. Motility showed the highest relationship with SCR in both time points and for both the Pearson and Spearman methods but was still less than 0.20. Parameters such as VAP-m, VCL-m, ALH-m, HSZ-m, VAP-v, VSL-v, VCL-v, ELO-v had correlation coefficients with SCR of similar magnitudes but in the negative direction in at least

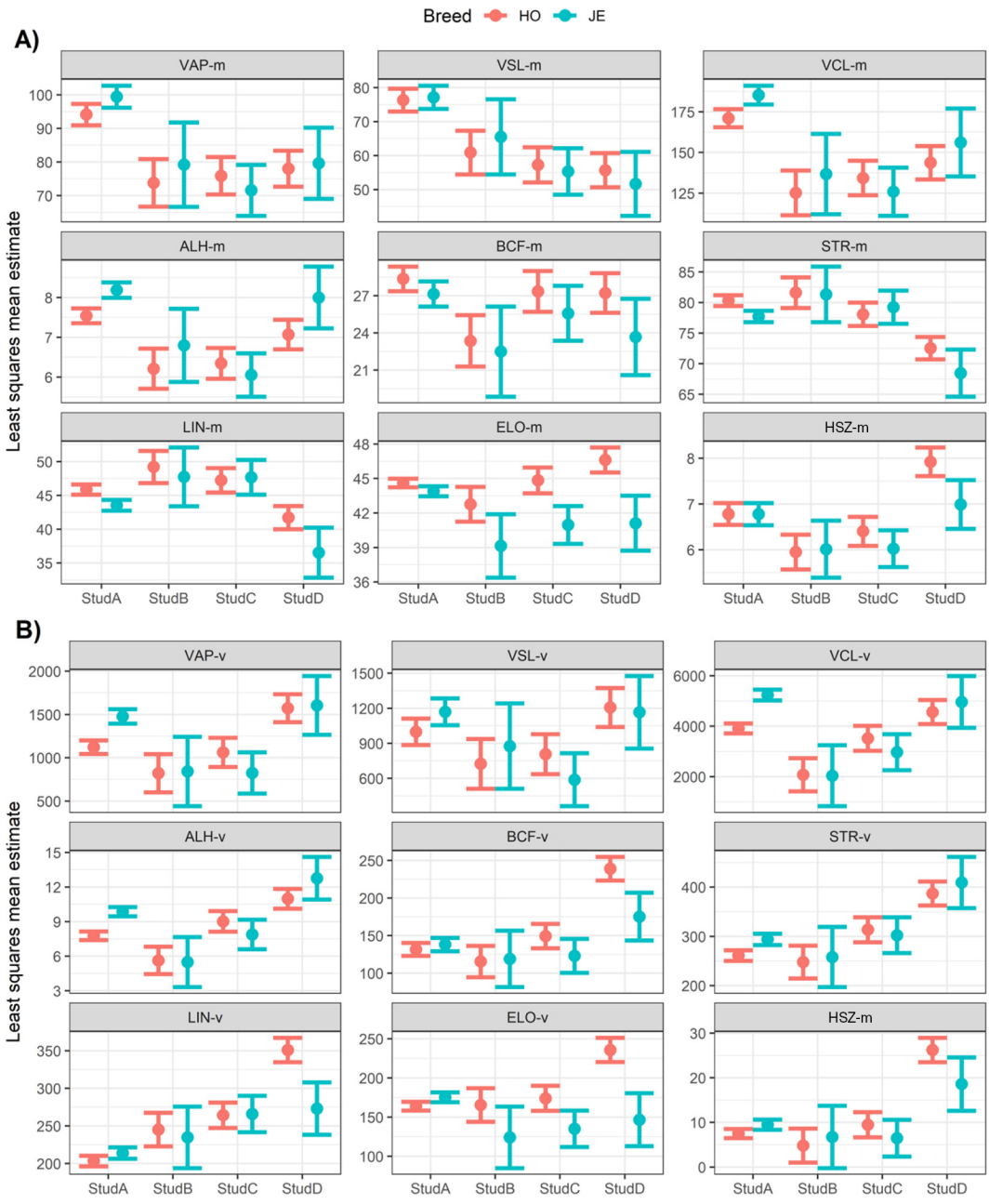

Figure 1. Least squares means and standard error (bars) estimates of the mean (A) and variance (B) of sperm kinematic and morphometric parameters of bulls by breed (HO = Holstein and JE = Jersey) and AI stud (A to D).

one of the tests and time points, while STR-m and LIN-m showed positive correlations.

Principal Components Analysis of CASA Parameters

Table 3 shows the variation explained and eigenvectors of the PC evaluated at T1 and T2. At each time, 4 PC largely explained the variation. Except for the mean and variance of ALH, all other kinematic parameters, as well as all morphological parameters, showed loadings on at least one of the specific PC. At both time points, we found that PC1 was mainly associated with variation in the mean and variance of sperm velocity parameters. The second PC (PC2) was associated with variation in the mean of velocity parameters in both time points, BCF-m in T2, STR-m in both time points, LIN-m in T2, STR-v

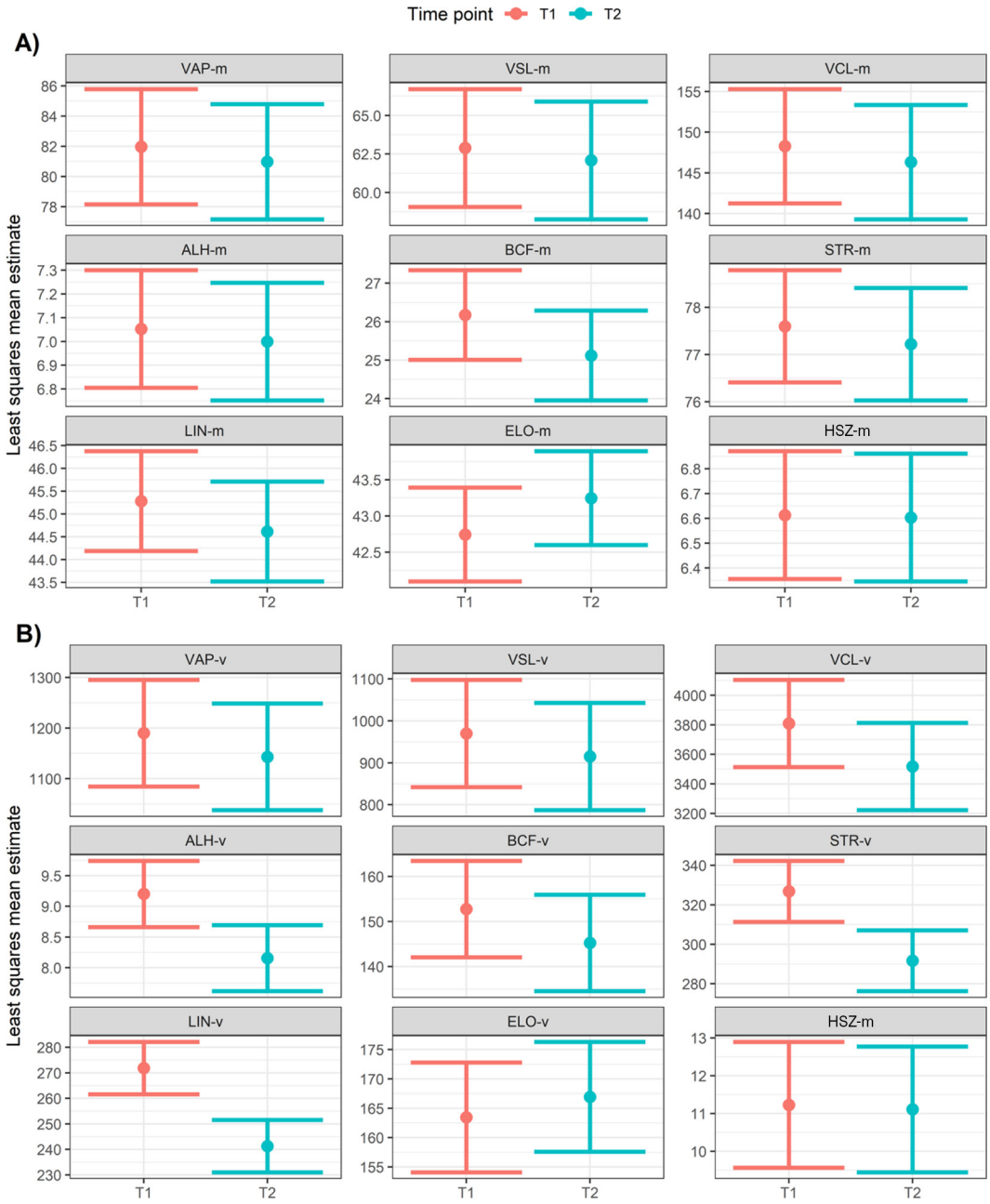

Figure 2. Least squares means and standard error (bars) estimates of the mean (A) and variance (B) of sperm kinematic and morphometric parameters of bulls by evaluation time point (0 and 30 min post-thawing).

in both time points, and LIN-v at T1. The sperm parameters associated with variation for PC3 were generally different between time points. At T1, PC3 was associated with variation for BCF-m, LIN-m, HSZ-m, BCF-v, and LIN-v. At T2, PC3 was also associated with BCF-v and LIN-v, in addition to VCL-m and HSZ-v. The fourth PC was strongly associated with variation in the mean and variance of sperm elongation at both time points (Supplemental Figure S3, see Notes).

Clustering of Bulls

Based upon the PCA, bulls were clustered for T1 and T2 (Supplemental Figure S3). Least squares means and CI for CASA parameters and SCR by cluster at T1 and T2 are shown in Tables 4 and 5, respectively.

The T1-BC1 cluster consisted of 180 bulls and had the highest SCR. In general, the bulls of this cluster also had the lowest values for CASA traits compared with

Table 3. Variation explained and eigenvectors of the first 4 principal components (PC) for bull sperm kinematic and morphometric parameters in 2 sample evaluation time points (T1 and T2)

	T1				T2			
Item	1-PC1	1-PC2	1-PC3	1-PC4	2-PC1	2-PC2	2-PC3	2-PC4
Explained variation (%)	38.92	19.48	14.73	8.47	35.29	24.33	15.47	8.09
Sperm parameter[1] (mean)								
VAP-m	−0.30	−0.30						
VSL-m		−0.43				−0.372		
VCL-m	−0.31				−0.312		−0.335	
ALH-m								
BCF-m			−0.44			−0.423		
STR-m		−0.38				−0.435		
LIN-m			−0.36			−0.434		
ELO-m				−0.68				0.705
HSZ-m			−0.33					
Sperm parameter[1] (variance)								
VAP-v	−0.33					−0.361		
VSL-v	−0.32				−0.335			
VCL-v	−0.33				−0.325			
ALH-v								
BCF-v			−0.36				0.34	
STR-v		0.39				0.309		
LIN-v		0.34	−0.35				0.461	
ELO-v				−0.59				0.578
HSZ-v							0.349	

[1]Sperm parameter: VAP = average path velocity, VSL = straight-line velocity, VCL = curvilinear velocity, ALH = amplitude of lateral head displacement, BCF = beat-cross frequency, STR = threshold straightness, LIN = linearity, ELO = elongation ratio, HSZ = head size.

the other clusters. Bulls assigned to the T1-BC1 group had sperm with smaller size (HSZ-m) and less elongated shape (ELO-m), which varied less in these characteristics (HSZ-v and ELO-v) compared with bulls assigned to other clusters. The cluster with the lowest average SCR was T1-BC3. This cluster's means for VAP, VSL, VCL, ALH, BCF, STR, LIN, and HSZ were larger or next to largest when compared with the other clusters. In particular, the percentage differences for T1-BC3 compared with T1-BC1 for VCL, ALH, and HSZ were 50.3%, 39.4%, and 16.8% larger, respectively. However, the within-cluster variances were not the largest, suggesting relative uniformity among sperm cells in this cluster.

Clusters T1-BC2 and T1-BC4 were ranked intermediate for fertility. Their sperm speed had higher variances, higher than T1-BC1, and their means were substantially lower than T1-BC3. Interestingly, for HSZ, their means were the same (T1-BC4) as T1-BC1 or the largest (T1-BC2) when compared with all clusters. The CASA variances for these 2 clusters were generally different. Cluster T1-BC2 had the highest variances among its CASA measures, suggesting within-group variability of sperm cells. Cluster T1-BC4 on the other hand had the second smallest variances of the majority of parameters (VAP, VSL, VCL, BCF, ELO, and HSZ), suggesting greater uniformity in this set of sperm cells. In part, the higher variances associated with T1-BC2 may have been due to the small number of bulls (n = 27 from 3 different studs) in the cluster. Among the 4 clusters at T1 T1-BC1, T1-BC3, and T1-BC4 had high levels of MOT and PMOT. However, T1-BC2 had the lowest MOT and PMOT and an intermediate SCR.

At T2 bulls were categorized into 2 clusters. It is possible to see in Supplemental Table S3 (see Notes) that, in general, T1-BC1 and T1-BC4 at T1 were combined to form T2-BC2, which was the cluster with the highest number of bulls at T2. In contrast, T1-BC2 and T1-BC3 were generally merged to form T2-BC1 at T2 with a smaller number of bulls. As a set of CASA parameters, T2-BC2 had lower values, as with T1-BC1, along with the highest SCR.

DISCUSSION

The CASA analyses provide an objective measurement of sperm motility and morphology, but they do not provide estimates of the subjective values of quality or fertilizing potential of a sperm sample. Consequently, when only CASA analyses are considered, total motility or progressive motility have been the sole means of estimating quality and fertility, and those relationships are not fully founded. For this reason, a model that can more fully and accurately explain the relationship between CASA values and fertilizing potential is needed. Although many studs also use flow cytometric and morphological evaluations to provide a comprehensive assessment of sperm quality, the experiments described here focused exclusively on the relationship of CASA and SCR. Therefore,

Table 4. Least squares means estimates (95% CI) for sperm kinematic and morphometric parameters, sample motility parameters, and sire conception rate (SCR) of bull clusters (BC) for sample evaluation time point 1 (T1)

	T1 bull cluster[1]			
Parameter[2]	T1-BC1 (n = 180)	T1-BC2 (n = 27)	T1-BC3 (n = 101)	T1-BC4 (n = 151)
Kinematic (mean)				
VAP-m (μm/s)	81.95 (80.69, 83.22)	94.12 (90.86, 97.37)	115.82 (114.14, 117.5)	89.98 (88.61, 91.36)
VSL-m (μm/s)	67.5 (66.47, 68.53)	67.45 (64.8, 70.1)	94.69 (93.32, 96.06)	66.37 (65.24, 67.49)
VCL-m (μm/s)	144.36 (141.6, 147.12)	168 (160.88, 175.11)	216.99 (213.31, 220.67)	170.13 (167.12, 173.14)
ALH-m (μm)	6.57 (6.45, 6.69)	7.41 (7.1, 7.72)	9.15 (8.99, 9.31)	7.86 (7.73, 7.99)
BCF-m (Hz)	28.56 (28.15, 28.98)	30.53 (29.46, 31.6)	30.47 (29.92, 31.02)	25.23 (24.77, 25.68)
STR-m (%)	82.29 (81.84, 82.74)	71.67 (70.5, 72.84)	80.96 (80.36, 81.57)	74.97 (74.47, 75.46)
LIN-m (%)	48.9 (48.43, 49.37)	42.41 (41.18, 43.63)	45.06 (44.43, 45.69)	41.55 (41.03, 42.07)
Kinematic (variance)				
VAP-v	885.47 (833.9, 937.05)	2,168.62 (2,035.46, 2,301.78)	1,578.58 (1,509.74, 1,647.43)	1,316.08 (1,259.77, 1,372.38)
VSL-v	766.37 (725.07, 807.67)	1,708.3 (1,601.67, 1,814.94)	1,492.49 (1,437.35, 1,547.62)	946.74 (901.65, 991.83)
VCL-v	3,149.91 (2,984.11, 3,315.71)	5,881.49 (5,453.41, 6,309.58)	5,779.05 (5,557.72, 6,000.38)	4,734.18 (4,553.16, 4,915.2)
ALH-v	7.2 (6.9, 7.5)	12.75 (11.98, 13.52)	10 (9.6, 10.4)	10.09 (9.76, 10.41)
BCF-v	125.33 (121.11, 129.56)	257.39 (246.48, 268.3)	148.82 (143.18, 154.46)	137.05 (132.44, 141.67)
STR-v	255.91 (247.76, 264.07)	426.85 (405.79, 447.91)	265.57 (254.69, 276.46)	355.83 (346.92, 364.73)
LIN-v	232.33 (226.39, 238.27)	358 (342.66, 373.34)	203.36 (195.43, 211.29)	237.28 (230.8, 243.77)
Morphometric (mean)				
ELO-m (%)	43.7 (43.32, 44.08)	46.81 (45.83, 47.8)	44.06 (43.55, 44.57)	44.13 (43.71, 44.54)
HSZ-m (μm^2)	6.42 (6.37, 6.48)	8.18 (8.03, 8.33)	7.5 (7.43, 7.58)	6.37 (6.3, 6.43)
Morphometric (variance)				
ELO-v	149.15 (143.37, 154.92)	248.73 (233.82, 263.64)	183.48 (175.77, 191.18)	170.19 (163.88, 176.49)
HSZ-v	7.56 (6.47, 8.64)	22.56 (19.76, 25.37)	10.05 (8.6, 11.5)	8.42 (7.24, 9.61)
Sample motility (mean)				
MOT (%)	54.21 (52.47, 55.95)	29.7 (25.22, 34.19)	48.75 (46.43, 51.07)	54.1 (52.2, 56)
PMOT (%)	37.95 (36.68, 39.22)	15.19 (11.9, 18.47)	37.12 (35.42, 38.82)	31.66 (30.27, 33.04)
Realized fertility (mean)				
SCR (−0+)	−0.07 (−0.39, 0.25)	−0.79 (−1.66, 0.08)	−1.29 (−1.73, −0.85)	−0.8 (−1.15, −0.44)

[1]T1-BC = bull cluster at T1.

[2]Parameters: VAP = average path velocity, VSL = straight-line velocity, VCL = curvilinear velocity, ALH = amplitude of lateral head displacement, BCF = beat-cross frequency, STR = threshold straightness, LIN = linearity, ELO = elongation ratio, HSZ = head size, MOT = percentage of motility sperm, PMOT = percentage of progressive motility sperm.

this research intended to develop a method that coupled CASA results with multivariate statistical analyses to predict the fertilizing potential of dairy bulls based on SCR breeding values.

Each mating is a binary trait, however, as the number of matings per bull increases this can be represented by some distribution (normal or otherwise) and individual animals can be combined into a distribution for all animals and their matings. The point is that an infertile bull will be a rare occurrence in such a distribution and located within the distribution's left tail. As such, assessing bull fertility involves measuring degrees or proportional differences, rather than making a binary assessment. Furthermore, as infertile bulls are culled through artificial or natural selection, most males evaluated will have fertility at some level. To determine whether CASA values can predict fertilizing potential, we integrated the outcomes of the CASA-derived data with AI data via SCR.

Although it has been reported that there may be a difference of more than 10% in conception rates between bulls of higher and lower fertility, as measured by SCR (Peñagaricano et al., 2012), greater than 93% of all bulls on the April 2024 SCR list are between +3 and −3 (CDCB, 2024). Clustering bulls by similarity of CASA parameters was shown to result in different average SCR levels within the distribution of this dataset, suggesting its utility in projecting bull fertility and illustrating fertility differences. The various CASA parameters for each cluster had relatively unique ranges, suggesting as a group that they might be used in further work on fertility prediction.

Reporting only a single motility value may lead to an emphasis on high motility, potentially causing improper prioritization of bulls with the highest motility or other CASA parameters (Blackburn et al., 2022). Therefore, it was crucial to analyze the CASA parameters from the perspective of their average but also through the lens of their variance. Furthermore, a more holistic view of the results was needed, interpreting them not from just one or a few characteristics, but rather by considering all CASA values as potential indicators of quality in combination with all the effects used to determine the SCR for a bull.

Least squares means and CI were computed for each bull cluster at T1 and T2. Exploring CASA LSM by bull cluster indicated that CASA parameters and SCR differed by cluster and this was also influenced by breed, stud,

Table 5. Least squares means estimates (95% CI) for sperm kinematic and morphometric parameters, sample motility parameters, and sire conception rate (SCR) of bull clusters (BC) for sample evaluation time point 2 (T2)

	T2 bull cluster[1]	
Parameter[2]	T2-BC1 (n = 129)	T2-BC2 (n = 330)
Kinematic (mean)		
VAP-m (μm/s)	107.35 (105.63, 109.06)	85.22 (84.15, 86.29)
VSL-m (μm/s)	87.14 (85.36, 88.92)	66.29 (65.18, 67.4)
VCL-m (μm/s)	192.93 (189.42, 196.44)	157.39 (155.2, 159.59)
ALH-m (μm)	8.09 (7.93, 8.25)	7.29 (7.19, 7.39)
BCF-m (Hz)	31.2 (30.52, 31.88)	25.42 (25, 25.85)
STR-m (%)	79.99 (79.01, 80.97)	78.19 (77.58, 78.81)
LIN-m (%)	46.19 (45.19, 47.2)	44.18 (43.55, 44.81)
Kinematic (variance)		
VAP-v	1,614.89 (1,562.14, 1,667.63)	1,038.93 (1,005.95, 1,071.9)
VSL-v	1,463.68 (1,421.38, 1,505.98)	790.13 (763.69, 816.58)
VCL-v	5,169.54 (4,973.86, 5,365.22)	3,681.42 (3,559.07, 3,803.77)
ALH-v	9.74 (9.33, 10.15)	7.4 (7.15, 7.66)
BCF-v	173.15 (166.49, 179.8)	119.02 (114.86, 123.18)
STR-v	279.09 (268.15, 290.02)	259.29 (252.46, 266.13)
LIN-v	223.63 (215.12, 232.14)	197.93 (192.61, 203.25)
Morphometric (mean)		
ELO-m (%)	45.9 (45.32, 46.48)	44.05 (43.69, 44.42)
HSZ-m (μm^2)	7.55 (7.45, 7.66)	6.41 (6.34, 6.48)
Morphometric (variance)		
ELO-v	205.51 (197.02, 214.01)	159.64 (154.33, 164.95)
HSZ-v	14.18 (12.49, 15.88)	7.35 (6.3, 8.41)
Sample motility (mean)		
MOT (%)	45.02 (42.64, 47.39)	52.92 (51.43, 54.41)
PMOT (%)	32.57 (30.75, 34.39)	33.84 (32.7, 34.98)
Realized fertility (mean)		
SCR (−0+)	−0.99 (−1.38, −0.59)	−0.46 (−0.7, −0.22)

[1]T2-BC = bull cluster at T2.

[2]Parameters: VAP = average path velocity, VSL = straight-line velocity, VCL = curvilinear velocity, ALH = amplitude of lateral head displacement, BCF = beat-cross frequency, STR = threshold straightness, LIN = linearity, ELO = elongation ratio, HSZ = head size, MOT = percentage of motility sperm, PMOT = percentage of progressive motility sperm.

their interactions, and the evaluation time. Our results indicated that having the highest average values for most kinetic parameters is not an indication of high fertility. Instead, higher fertility was observed with many lower CASA measures (T1-BC1 and T2-BC2) and the inverse was also true with bulls in the T1-BC3 cluster, which had the lowest fertility and tended to have the highest CASA measurements. This suggests that the dogma that the fastest speeds or the highest percentages of motile (total and progressive) sperm, for example, do not translate into greater fertility. This perspective revealed that lower variation in sperm kinetic and morphometric parameters, along with higher percentages of total and progressively motility sperm, may have a stronger relationship with bull AI fertility than any single CASA measurement. Furthermore, we observed that higher sperm velocity averages do not increase fertilization chances, which contradicts several findings that state that faster bull sperm have greater fertilization potential (Nagy et al., 2015; Ibanescu et al., 2020; Donnellan et al., 2022).

It is also critical to observe that the dataset that was used was substantially different from most other analyses. Often studies used between 10 and 24 males to identify the relationship between sperm characteristics and fertility (Nagy et al., 2015; Ibanescu et al., 2020; Donnellan et al., 2022), while our work used 459 bulls, likely increasing the diversity of the population studied. Furthermore, the bulls used in this study exhibited a greater range of CASA values than those in other research. We believe the dataset used in this study initiates exploration into the dynamics of CASA parameters, for example, VCL. Our results suggest an optimal range of values for each trait can yield acceptable fertility, but once outside that range of values (either higher or lower) fertility decreases. Furthermore, the optimal range for a trait could be affected by the value of other traits.

In general, what can be observed in our work was that bulls that possessed sperm exhibiting lower velocity, size (Gravance et al., 1998; Guthauser et al., 2011), and ALH (Gogol and Trzcińska, 2022), BCF, greater straightness and linearity, and lower variability in terms of kinematic and morphometric parameters may have an advantage in terms of fertility. These characteristics were all observed in the bulls of group T1-BC1 and partially observed in

the bulls of group T2-BC2, which presented higher SCR in T1 and T2, respectively. However, for these bulls to effectively have higher fertility, these characteristics need to be accompanied by an acceptable percentage of motile sperm with progressive movement, as exemplified in T1-BC1 (54.21% and 37.95%) and T2-BC-2 (52.92% and 33.84%). Conversely, the other groups of bulls in T1 or T2 that presented higher values of velocity, ALH, BCF, and HSZ of the spermatozoa, and greater variances in the kinematic and morphometric parameters, in addition to lower percentages of total and progressive motility in semen, showed lower SCR. High values of these parameters can characterize a hyperactive movement of the spermatozoa (Yanagimachi, 1994), which ideally should not be present immediately after thawing when a bull sperm sample is intended for AI. Instead, sperm hyperactivation should only manifest in the uterine environment (Ho and Suarez, 2001). Thus, the bulls from the groups with lower SCR likely had semen with hyperactivated spermatozoa and with a lower probability of fertilization. The results confirm that the evaluation of bull fertility via frozen-thawed sperm must consider multiple CASA values, rather than single CASA values, to make the model more robust. Moreover, it is known that these are not the only characteristics that define healthy and fertile sperm function and therefore future analyses must consider cell physiology, evaluated via flow cytometry, and morphological characteristics. In doing so, we may be able to account for a greater proportion of the variance when predicting fertility.

CONCLUSIONS

This study used statistical approaches to evaluate CASA-generated sperm parameters. Clustering individuals based upon CASA parameters could result in clusters comprised of unique sets (where each set has specific ranges of values) of CASA parameters and differences between the clusters and their sets of CASA parameters corresponded to different SCR. This suggests to effectively evaluate sperm for fertilizing potential, unique sets of parameters should be used, with each set leading to different fertility levels. The range of CASA parameters within a set can influence the males' fertility level. This suggests clustering could accommodate the kind of distributional properties observed in comparing fertility levels of bulls. This is a departure from past approaches used with CASA in the hopes of building prediction equations for fertility. Although BC1-T1 had the highest SCR, the other clusters are not infertile, but simply had lower levels of fertility. This suggests a plasticity among combinations of CASA parameters. We speculate this plasticity has made fertility prediction difficult when using CASA. The formation of clusters, each containing limited ranges for CASA parameters, is suggestive that sets of parameters exist that, when combined, result in varying levels of SCR for the respective cluster. Further exploration of this question is needed to define the properties of identified clusters and their sets of CASA parameters.

NOTES

The authors thank the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico–CNPq, Brasilia, Brazil), process 401006/2022-0, for financial support for H. C. Azevedo. Supplemental material for this article is available at https://doi.org/10.6084/m9.figshare.27190482.v2. The USDA Agricultural Research Service is an equal opportunity/affirmative action employer. The mention of companies or commercial products does not imply recommendation or endorsement by the USDA over others not mentioned. The USDA neither guarantees nor warrants the standard of any product mentioned. Product names are mentioned solely to report factually on available data and provide specific information. This work did not require ethical approval because it did not use animals but only semen doses from commercial AI companies (studs) for its execution. The authors have not stated any conflicts of interest.

Nonstandard abbreviations used: ALH = amplitude of lateral head displacement; BC = bull cluster; BCF = beat-cross frequency; CASA = computer-assisted sperm analysis; ELO = elongation ratio; HO = Holstein; HSZ = head size; JE = Jersey; LIN = linearity; m = mean; MOT = percentage of motile sperm; PC = principal component; PCA = principal components analysis; PMOT = percentage of progressively motile sperm; SCR = sire conception rate; STR = threshold straightness; T1 = immediately after thawing evaluation time point; T2 = 30 minutes after the first evaluation time points; v = variance; VAP = average path velocity; VCL = curvilinear velocity; VSL = straight-line velocity.

REFERENCES

Alquézar-Baeta, C., S. Gimeno-Martos, S. Miguel-Jiménez, P. Santolaria, J. Yániz, I. Palacín, A. Casao, J. Á. Cebrián-Pérez, T. Muiño-Blanco, and R. Pérez-Pé. 2019. OpenCASA: A new open-source and scalable tool for sperm quality analysis. PLOS Comput. Biol. 15:e1006691. https://doi.org/10.1371/journal.pcbi.1006691.

Amann, R. P., and D. Waberski. 2014. Computer-assisted sperm analysis (CASA): Capabilities and potential developments. Theriogenology 81:5–17.e3. https://doi.org/10.1016/j.theriogenology.2013.09.004.

Barth, A. D. 1989. Evaluation of frozen bovine semen. Pages 92–99 in Proc. Soc. Theriogenol. Society for Theriogenology.

Bates, D., M. Mächler, B. M. Bolker, and S. C. Walker. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67:1–48. https://doi.org/10.18637/jss.v067.i01.

Blackburn, H., L. Torres, Y. Liu, and T. R. Tiersch. 2022. The need for a framework addressing the temporal aspects of fish sperm motility leading to community-level standardization. Zebrafish 19:119–130. https://doi.org/10.1089/zeb.2022.0006.

Castro, L. S., T. R. S. Hamilton, C. M. Mendes, M. Nichi, V. H. Barnabe, J. A. Visintin, and M. E. O. A. Assumpção. 2016. Sperm cryodamage occurs after rapid freezing phase: Flow cytometry approach and antioxidant enzymes activity at different stages of cryopreservation. J. Anim. Sci. Biotechnol. 7:17. https://doi.org/10.1186/s40104-016-0076-x.

CDCB. 2024. Council on Dairy Cattle Breeding. Accessed Aug. 14, 2024. https://www.uscdcb.com/.

Donnellan, E. M., P. Lonergan, K. G. Meade, and S. Fair. 2022. An ex-vivo assessment of differential sperm transport in the female reproductive tract between high and low fertility bulls. Theriogenology 181:42–49. https://doi.org/10.1016/j.theriogenology.2022.01.011.

Ehlers, J., M. Behr, H. Bollwein, M. Beyerbach, and D. Waberski. 2011. Standardization of computer-assisted semen analysis using an e-learning application. Theriogenology 76:448–454. https://doi.org/10.1016/j.theriogenology.2011.02.021.

Farrell, P., G. Presicce, C. Brockett, and R. Foote. 1998. Quantification of bull sperm characteristics measured by computer-assisted sperm analysis (CASA) and the relationship to fertility. Theriogenology 49:871–879. https://doi.org/10.1016/S0093-691X(98)00036-3.

Gogol, P., and M. Trzcińska. 2022. Relationship between quality parameters and fertilizing ability of cryopreserved sexed bull sperm. Ann. Anim. Sci. 22:1257–1263. https://doi.org/10.2478/aoas-2022-0046.

Gravance, C. G., R. Vishwanath, C. Pitt, D. L. Garner, and P. J. Casey. 1998. Effects of cryopreservation on bull sperm head morphometry. J. Androl. 19:704–709. https://doi.org/10.1002/j.1939-4640.1998.tb02079.x.

Guthauser, B., M. Albert, F. Ferfouri, P. F. Ray, G. Rabiey, J. Selva, and F. Vialard. 2011. Inverse correlation between chromatin condensation and sperm head size in a case of enlarged sperm heads. Reprod. Biomed. Online 23:711–716. https://doi.org/10.1016/j.rbmo.2011.07.011.

Ho, H. C., and S. S. Suarez. 2001. Hyperactivation of mammalian spermatozoa: Function and regulation. Reproduction 122:519–526. https://doi.org/10.1530/rep.0.1220519.

Ibanescu, I., M. Siuda, and H. Bollwein. 2020. Motile sperm subpopulations in bull semen using different clustering approaches—Associations with flow cytometric sperm characteristics and fertility. Anim. Reprod. Sci. 215:106329. https://doi.org/10.1016/j.anireprosci.2020.106329.

Jeong, J., and S. C. Chong. 2020. Adaptation to mean and variance: Interrelationships between mean and variance representations in orientation perception. Vision Res. 167:46–53. https://doi.org/10.1016/j.visres.2020.01.002.

Kasimanickam, R., R. L. Nebel, I. D. Peeler, W. L. Silvia, K. T. Wolf, A. J. McAllister, and B. G. Cassell. 2006. Breed differences in competitive indices of Holstein and Jersey bulls and their association with sperm DNA fragmentation index and plasma membrane integrity. Theriogenology 66:1307–1315. https://doi.org/10.1016/j.theriogenology.2006.04.025.

Lenth, R. V. 2023. Package emmeans: Estimated Marginal Means, aka Least-Squares Means. Version 1.8.8. Accessed Sep. 13, 2023. https://rvlenth.github.io/emmeans/.

Li, Y., D. Kalo, Y. Zeron, and Z. Roth. 2016. Progressive motility—A potential predictive parameter for semen fertilization capacity in bovines. Zygote 24:70–82. https://doi.org/10.1017/S0967199414000720.

Nagy, Á., T. Polichronopoulos, A. Gáspárdy, L. Solti, and S. Cseh. 2015. Correlation between bull fertility and sperm cell velocity parameters generated by computer-assisted semen analysis. Acta Vet. Hung. 63:370–381. https://doi.org/10.1556/004.2015.035.

Norman, H. D., J. L. Hutchison, and J. R. Wright. 2008. Sire Conception Rate: New National AI Bull Fertility Evaluation. Accessed Feb. 20, 2024. http://aipl.arsusda.gov/reference/arr-scr1.htm.

Park, M. J., M. Y. Lim, H. J. Park, and N. C. Park. 2021. Accuracy comparison study of new smartphone-based semen analyzer versus laboratory sperm quality analyzer. Investig. Clin. Urol. 62:672–680. https://doi.org/10.4111/icu.20210266.

Peñagaricano, F., K. A. Weigel, and H. Khatib. 2012. Genome-wide association study identifies candidate markers for bull fertility in Holstein dairy cattle. Anim. Genet. 43(Suppl. 1):65–71. https://doi.org/10.1111/j.1365-2052.2012.02350.x.

Purdy, P. H., and J. K. Graham. 2004. Effect of cholesterol-loaded cyclodextrin on the cryosurvival of bull sperm. Cryobiology 48:36–45. https://doi.org/10.1016/j.cryobiol.2003.12.001.

R Core Team. 2023. R: A Language and Environment for Statistical Computing. Version 4.3.1. R Foundation for Statistical Computing. Vienna, Austria.

Rodriguez, M. Z., C. H. Comin, D. Casanova, O. M. Bruno, D. R. Amancio, L. Da F. Costa, and F. A. Rodrigues. 2019. Clustering algorithms: A comparative approach. PLoS One 14:e0210236. https://doi.org/10.1371/journal.pone.0210236.

Silva, C. S., E. V. da Costa-e-Silva, M. A. N. Dode, A. T. Mendes Cunha, W. R. Garcia, B. F. B. Sampaio, J. C. Borges-Silva, F. E. M. Vaz, K. Kerns, P. Sutovsky, and E. Nogueira. 2023. Semen quality of Nellore and Angus bulls classified by fertility indices and relations with field fertility in fixed-time artificial insemination. Theriogenology 212:148–156. https://doi.org/10.1016/j.theriogenology.2023.09.001.

Verstegen, J., M. Iguer-Ouada, and K. Onclin. 2002. Computer assisted semen analyzers in andrology research and veterinary practice. Theriogenology 57:149–179. https://doi.org/10.1016/S0093-691X(01)00664-1.

Yanagimachi, R. 1994. Mammalian fertilization. Pages 189–317 in The Physiology of Reproduction, E. Knobil and J. Neill, ed. Raven Press, New York.

Zoca, S. M., T. W. Geary, A. L. Zezeski, K. C. Kerns, J. C. Dalton, B. R. Harstine, M. D. Utt, R. A. Cushman, J. A. Walker, and G. A. Perry. 2023. Bull field fertility differences can be estimated with in vitro sperm capacitation and flow cytometry. Front. Anim. Sci. 4:1180975. https://doi.org/10.3389/fanim.2023.1180975.

ORCIDS

H. C. Azevedo, https://orcid.org/0000-0002-0187-9227
H. D. Blackburn, https://orcid.org/0000-0003-0575-1594
E. A. Lozada-Soto, https://orcid.org/0000-0001-8381-7988
S. F. Spiller, https://orcid.org/0009-0008-5061-885X
P. H. Purdy, https://orcid.org/0000-0001-7456-2629