Sperm nuclear texture assessment: a fertility prediction tool

Introduction

Sperm nuclear morphometry quantifies properties of the nucleus and allows comparison between semen samples. Chromatin packaging is one of the properties reflected by nuclear measures, and has been related to fertility as correct packaging is essential for not only fertilization but also successful early embryonic development. One such morphometric measure is texture. Nuclear texture is thought to reflect the structure of the chromatin within sperm and the degree and pattern of packaging, properties vital to successful fertilization and proper embryonic development. Aberrations in sperm chromatin packaging have been linked to failure in the fertilizing ability of semen samples (Engh et al., 1993; Evenson et al., 2002; Filatov et al., 1999; Ostermeier, 2000; Sailer et al., 1996).

Advancements in image analysis, aided by computer software improvements and the development of consistent and precise fluorescent nanoparticles, continually provide new ways to assess fluorescent images and describe nuclear parameters. Image deconvolution, the process that estimates and removes blur from images is one such example. Deconvolution is achieved by imaging spherical fluorescent nanoparticles that serve as a fluorescent point source, from which a point spread function (PSF) is developed. The PSF identifies the fundamental unit of an image by defining the diffraction pattern created by imperfections in the lens of an imaging system. Iterative algorithms utilizing the PSF remove blur from fluorescent images acquired under the same lens and magnifiers. The result is a reconstructed image that should more accurately reflect the imaged object.

Deconvolved images of chromatin bound dye in sperm allow differences in the texture to be more pronounced. The present study reports the relationship between fertility and sperm nuclear texture with the application of deconvolution as an aid in image analysis. The ultimate goal is to establish a new tool for identifying low fertility animals or semen samples. The exclusion of such low fertility samples from the breeding stock or for selection of high fertility semen samples for use in In vitro fertilization would benefit advanced reproductive technologies.

Materials and Methods

Semen Samples

Frozen semen samples from 107 bulls, from various dairy farms throughout the U.S., were received from Alta Genetics. The bulls were of varied fertility with Estimated Relative Conception Rates (ERCR) ranging from -13.08 to 6.87 expressed as the percent deviation of conception rate from the average conception of all bulls, calculated from service records ranging from 400 to 11,760 inseminations per individual. All bulls were a minimum of one standard deviation (SD) above or below the mean, to allow the designation of high and low fertility groups.

Media

The following media were used for slide preparation: 2.9% Sodium Citrate Buffer (2.9 g sodium citrate dehydrate, 90 mL distilled water, pH 7.4, final volume adjusted to 100 mL); Hepes buffered saline (0.238 g Hepes free acid, 0.9 g NaCl, 90 mL distilled water, pH to 7.4, final volume adjusted to 100 mL with distilled water); DABCO mounting media (25 mg 1,4-Diazabicycolo[2.2.2] octane Triethylenediamine as anti-fade reagent, 100 µL Hepes buffered saline, mixed until dissolved, 900 µL glycerol, stored in a foil-wrapped tube); Paraformaldehyde (PFA) stock solution (4%, 4 g paraformaldehyde, 50 mL water, mixed, NaOH pellets added while mixing to dissolve PFA, 0.238 g Hepes, pH adjusted to 7.4, distilled water

added to 100 mL); Parrish citrate fixative (10 mL PFA stock, filled to 100 mL with 2.9% sodium citrate buffer); Hoechst 33342 (Life Technologies) stain solution (5mg/mL in water, made fresh daily).

Slide Preparation

For each bull a 0.5 mL straw was thawed in a water bath at 37 ℃ for 30 sec, ejected into a 2.0 mL centrifuge tube, to which 1 mL Parrish fixative with 3 mg/mL BSA was added. Samples were stored at 4℃ overnight to allow for complete fixation. Concentration was adjusted to 40 million sperm/mL and 500 µL was placed into a 2.0 mL centrifuge tube. Sperm were stained by adding 2.5 µL Hoechst 33342, mixing via vortex then incubated at 37℃ for 30 min.

Following the staining period, excess dye was removed by centrifugation at 6000 g for 15 sec, followed by aspiration to pellet. The pellet was re-suspended in 750 µL Parrish citrate fixative, and centrifugation, aspiration and re-suspension in fixative was repeated. The same cycle of two washes was repeated again but with re-suspension into PBS. A final cycle of centrifugation and aspiration was followed by re-suspension into 500 µL water, to the original concentration of 40 million spermatozoa/mL. Each sample was then mixed via vortex and 10 µL was spread onto a clean slide and allowed to dry on a slide warmer for a minimum of 2 hours, while being protected from light with tin foil. DABCO mounting media was applied to semen on slide (3.5uL) and then covered with an 18x18x1 mm coverslip and sealed with clear fingernail polish.

Image Acquisition

Slides were imaged with a QIClick monochrome camera operating in 8-bit mode fitted onto a Nikon Microphot fluorescence microscope (excitation filter 365 ± 20 nm, dichromatic mirror 400 nm, emission > 400 nm) excited with a Hg epifluorescent bulb. Image acquisition occurred under a 40x objective and 1.25 X magnifier with a neutral density filter in place and a 62.5 ms exposure length. Each field imaged was captured first as a phase contrast image, and then as fluorescent image.

Image Analysis

Point Spread Function Development and Deconvolution

A PS-Speck Microscope Point Source Kit (Life Technologies, Eugene, OR) containing fluorescent microspheres 0.175 ±0.005 µm in size with emission/excitation wavelengths 360/440 nm was used. Microspheres in suspension were diluted 1:50 in water, 10 µL of which was dried onto slides coated with Poly-L-lysine. Slide coating was created by applying 100 µL of 1 mg/mL Poly-L-lysine to slides for 5 min, after which excess was aspirated from slide surface and the remainder was dried. Once microsphere solution was dried to slide, mountant and coverslip were applied as for semen above.

The prepared slide was then imaged on the Nikon Microphot imaging system described above, with the exception of exposure time being increased to 500 ms. Eight images were acquired and one microsphere was selected from each to be cut from the larger image containing numerous microspheres. The cut spheres were then pasted and centered into new images of black backgrounds, sized to be the same as those of sperm (1392 x 1040 pixels). All image manipulations were performed with ImageJ v.1.49o software. The plugin image calculator was used to average the eight images together and the resulting image of a centered dot was used as the PSF for the deconvolution process.

The image sets of all 107 bull semen samples were copied, and deconvolution was performed on the fluorescent images. ImageJ plugin “2D Parallel Iterative Deconvolution” (Wendykier et al., 2009) was run in Wien Filter Preconditioned Landweber (WPL) mode which is an iterative method of blurring and

un-blurring by application of algorithms to image intensity measurements. Other settings within the plugin were selected as follows: boundary as reflexive, resizing as auto, output same as source, precision type as single, iterations as 5, and max number of threads as 8. The result for each image is a second image with the blur removed, showing an increased contrast across the sperm nuclei (see Figure 1).

ImageJ macros developed by the Parrish lab were used to identify the sperm perimeters and define them as the regions of interest to be analyzed. Perimeter outlines identifying sperm nuclei were determined within the original fluorescent image. These outlines were used to define areas of texture analysis on both the original fluorescent images and also the deconvolved image sets.

Texture Analysis

Sperm nuclear texture was measured using the plugin "GLCM Texture Analyzer" (Cabrera, 2006) for ImageJ. The plugin computes four texture parameters (angular second moment (ASM), contrast, inverse difference moment (IDM) and entropy) as described by Haralick et al. (1973) and one texture parameter (correlation) as defined by Walker et al. (1995). The underlying basis for analysis is the construction of a grey level co-occurrence matrix of conditional probabilities to which algorithms are applied (Table 1), thus deriving the texture parameters. These parameters summarize important information about the structural arrangement of surfaces by discerning likelihoods that pixels have the same or different grey-level values as their neighbors and the distances between pixel pairs of equal intensity.

The plugin was applied to both the deconvolved and original fluorescent images separately. It was limited to working with rectangular shapes and presumably chose the largest rectangular area within the identified sperm nucleus to analyze. The plugin was run at four degree angles (0, 90, 180, and 270) at a one pixel step interval. . Accordingly, the output consisted of the five texture parameters, at each of four degree angles for each sperm, in each image type. The mean value across all four degree angles was calculated for each parameter, for each sperm per image. One-hundred sperm were randomly selected for analysis from each semen sample. The mean between the 100 sperm was then calculated for each of five variables within bulls and for each image type the same 100 sperm were used.

Statistical Analysis

Analysis was performed in Statistical Analysis Software (SAS) version 9.3 unless otherwise specified. There were a total of 107 bulls, 53 of which were classified as high fertility, and 54 low. Data from the deconvolved and unaltered image sets were considered separately. A one-way analysis of variation (ANOVA) was performed comparing the five texture parameters individually between high and low fertility groups, followed by a multivariate analysis of variation (MANOVA) assessing the difference in all five texture values together between high and low fertility groups.

Discriminant analysis was performed to test the predictive power of using combined texture values for a given semen sample to correctly classify fertility. Preliminary observations of the data suggested that the variance within samples for each texture parameter were also different between fertility groups. The standard error of the means were normalized to z-scores to be considered in discriminant models along with the texture variables.

Discriminant analysis relies on the assumption of normality. Outliers were present in the dataset from deconvolved images. Outliers were identified with box and whisker plots (Figure 2) and a third dataset was created consisting of a subset original deconvolved dataset but with the 10 outliers removed. Discriminant analysis was performed on this as well.

ROC curves were made in conjunction with the discriminant analysis of all 5 texture variables plus normalized standard errors. These graphs were made in R (R Core Team, 2015), version 3.2.1, and the area under the curve (AUC) was calculated.

Results

Four of the five texture values from the unaltered fluorescent images were different between fertility groups ($p \leq 0.01$), and one from the deconvolved fluorescent images differed between fertility groups ($p = 0.025$) (Table 2.) The normalized standard errors were all different for the original image dataset and two of the five were different from the deconvolved images. The five texture parameters considered in aggregate were different between fertility groups ($p=0.01$) for each type of image: unaltered or deconvolved.

Angle second moment was the only texture value that was not independently related to fertility for either dataset. The angle second moment describes the uniformity of texture, i.e. the frequency of pixel pair repetitions. The three parameters contrast, correlation and entropy were different by fertility for the original fluorescent image dataset only. Contrast measures the difference between the highest and lowest pixels in a set, correlation measures the linear dependence between intensity values and entropy measures the disorder or complexity within an image.

Inverse different moment (IDM) was different between fertility groups for both the deconvolved and unaltered image sets, but in opposite directions. IDM describes the homogeneity of an image, a higher value reflecting more similarity between pixels. Within unaltered images lower fertility bulls show less nuclear homogeneity (more intensity difference) across the sperm nuclei than do high fertility bulls. The deconvolved dataset indicated a lower IDM value for the high fertility group.

Discriminant analysis used selected variables to derive a quadratic function that classified samples into high or low fertility groups. We used numerous models, dropping and adding texture parameters in various combinations for each image type in an effort to increase correct classification percentages especially into the low fertility group, and decrease in the overall error rate. Emphasis was also placed on retention of high percentages correct in cross-validation, as the initial classification is the most optimistic outcome for the function. The values showing the lowest correlation coefficients were dropped in some models and the subset of data with outliers removed was also assessed.

Discriminant analysis using all five texture values for each the unaltered and deconvolved datasets remained the best models for correctly identifying low fertility samples and retaining correctness in cross validation. The function maximizing correct classification from the deconvolved dataset correctly classified 83% of the low fertility samples and 45 % of the high fertility. Values from the original images gave a function correctly classifying 87% of the low and 49% of the high fertility bulls (See Table 3). Cross validation analysis correctly identified 78 and 72% low fertility and 36 and 30% of the high fertility for the deconvolved and unaltered dataset based functions respectively. These results are summarized with receiver operator curves (Figure 3).

Both analysis put a total of 69% of bulls into the low fertility group and 31% into the high, where the actual sample group had 50% in each. The function derived from the unaltered image dataset, however has a slightly lower error rate of 32% as opposed to 36% from the deconvolved derived function. The canonical structures show the correlations between each independent texture value and the function used for discrimination respective for each dataset (Table 4). Generally the correlation coefficients are higher for the original image derived texture values than for the deconvolved set. This was expected as the

univariate analysis results indicated a stronger relationship for each of the texture parameters to fertility for this dataset.

Removal of outliers did not result in higher percentages correct; 59% correct for high fertility and 75% for low. They did give greater consistency between the initial data classification and cross validation with 57% high and 75% low, though even the cross validated percentages were not as high as those achieved with other models. The addition of normalized standard errors into discrimination model increased the percent correct for the deconvolved dataset only. The increase, to 87% correctly identified low fertility bulls did not hold in cross-validation as the percentage dropped to 70%, lower than the other models. No other model explored resulted in increased percentages correct that resulted in cross-validation percentages higher than the models derived from five textures of original or deconvolved images.

Discussion

It was surprising that more textural features related to fertility from the unaltered images than from the deconvolved images. This may be a result of our deconvolution process removing too much fluorescence, resulting in a loss of information from the image. We chose an iterative process that applied the PSF and removed blur 5 successive times. It is possible that fewer iterations could still remove the blur but leave more intensity that is reflective of the chromatin, and therefore related to fertility.

Nuclear texture of sperm and fertility are established as related by both the univariate analysis performed on the original fluorescent images and multivariate and discriminant analysis of both image types. Overall the data indicate that the nuclei of sperm in low fertility samples have greater degrees of contrast and disorder while the nuclei of high fertility samples tend to have higher degrees of linear dependence and possibly homogeneity. Homogeneity responded differently between fertility groups depending on whether the image was deconvolved or unaltered.

The higher degree of homogeneity in the high fertility group, as seen for the original image dataset, agrees with the relationship between other texture variables to fertility: high fertility showing high degrees of linear dependence, and low fertility with higher values of entropy and contrast. The fact that the homogeneity relates differently to fertility for each dataset and that 4 of 5 individual texture parameters don't relate to fertility for the deconvolved dataset, but all 5 in aggregate do, indicates a complex relationship between fertility and texture descriptors. This fact remains true in the discriminant analysis, where the canonical correlations were not as high for texture values for the deconvolved dataset as for the original images, and yet the predictive power was roughly equal to that of the originals.

The values determined via texture analysis were found to be especially good at classifying low fertility semen samples. This suggests that the textural properties of sperm with abnormal chromatin features, conferring lower fertility, have more impact on texture. An example could be the case of a vacuole in the sperm head, showing up as a dark spot in the image, resulting in a high impact on texture values. General observation of the images tells us that vacuoles are uncommon, however there are more subtle variations in intensity across sperm. Texture analysis quantifies those differences and indicates they can be used to identify low fertility semen samples.

Conclusion

The findings herein report a moderate relationship between nuclear texture and fertility as found by image analysis. This preliminary study has also generated ideas for improvement upon the process that may increase the predictive ability as a fertility assessment technology. The texture analyzer plugin used

operates on rectangular ROI's only. A modified version of the plugin is now available, GLCM Texture Too (Cornish, 2007) that works on irregular ROI's. Areas for improvement include the use of this modified plugin and a trial of less deconvolution iterations in process of blur removal.

Texture analysis is a useful addition to fertility assessment and prediction, especially in the identification of low fertility animals. In domestic livestock species, where the majority of animals are in the high fertility group, but low fertility animals can be a significant source of production loss, such a model could be very valuable. Furthermore, nuclear texture analysis reflects differences in the chromatin of sperm, which is a non-compensable trait to fertility. Non-compensable traits to fertility may result in errors that occur during fertilization or embryo development, resulting in missed female cycles, an additional cost to production. In humans the value of identifying low fertility samples may aid in the selection of semen samples for use in advanced reproductive technologies such as In vitro fertilization, maximizing the chance for success thereby minimizing cost.

References

Cabrera, E. J. (2006). GLCM Texture Analyzer (Version v0.4). Retrieved from http://rsb.info.nih.gov/ij/plugins/texture.html

Cornish, C. T. (2007). GLCM Texture Too (Version v.008) [ImageJ Plugin]. ImageJ.

Engh, E., Scholberg, A., Clausen, O. P., & Purvis, K. (1993). DNA flow cytometry of sperm from normozoospermic men in barren couples and men of proven fertility. *Int J Fertil Menopausal Studd, 38*(5), 305-310.

Evenson, D. P., Larson, K. L., & Jost, L. K. (2002). Sperm chromatin structure assay: its clinical use for detecting sperm DNA fragmentation in male infertility and comparisons with other techniques. *J Androl, 23*(1), 25-43.

Filatov, M. V., Semenova, E. V., Vorob`eva, O. A., Leont`eva, O. A., & Drobchenko, E. A. (1999). Relationship between abnormal sperm chromatin packing and IVF results. *Molecular Human Reproduction, 5*(9), 825-830.

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems Man and Cybernetics, SMC-3*(6), 610-621.

Ostermeier, G. C. (2000). *Sperm nuclear morphology and its relationship with bull fertility.* (PhD Dissertation), University of Wisconsin-Madison.

R Core Team, A. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Sailer, B. L., Jost, L. K., & Evenson, D. P. (1996). Bull sperm head morphometery related to abnormal chromatin structure and fertility. *Cytometry*(24), 167-173.

Walker, R. F., Jackway, P., & Longstaff, I. D. (1995). *Improving Co-occurrence Matrix Feature Discrimination.* Paper presented at the Digital Image Computing: Techniques and Applications.

Wendykier, P., Nagy, J. G., Chung, J., Palmer, K., Perrone, L., Wright, R., & Dougherty, R. (2009). 2D Parallel Iterative Deconvolution.

Figure 1.

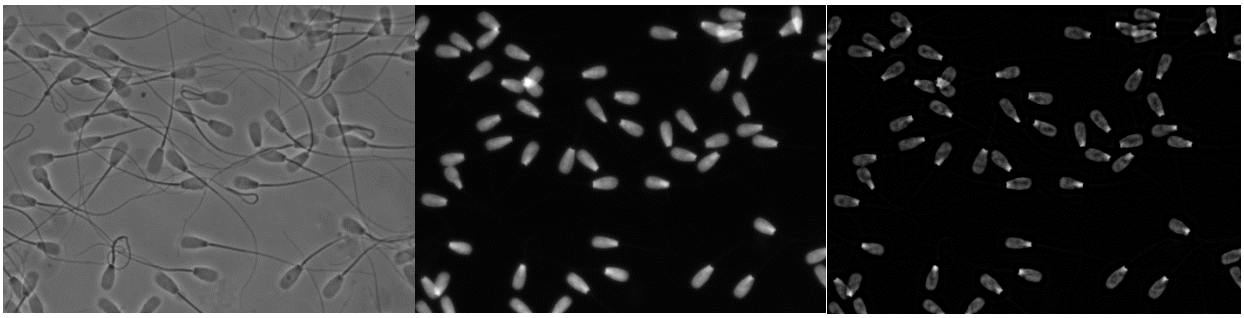


Image sequence showing (from left to right) the phase contrast image, fluorescent image, and the deconvolved fluorescent image. Texture analysis was performed on the deconvolved fluorescent image.

Figure 2.

Distribution of texture values by fertility groups.

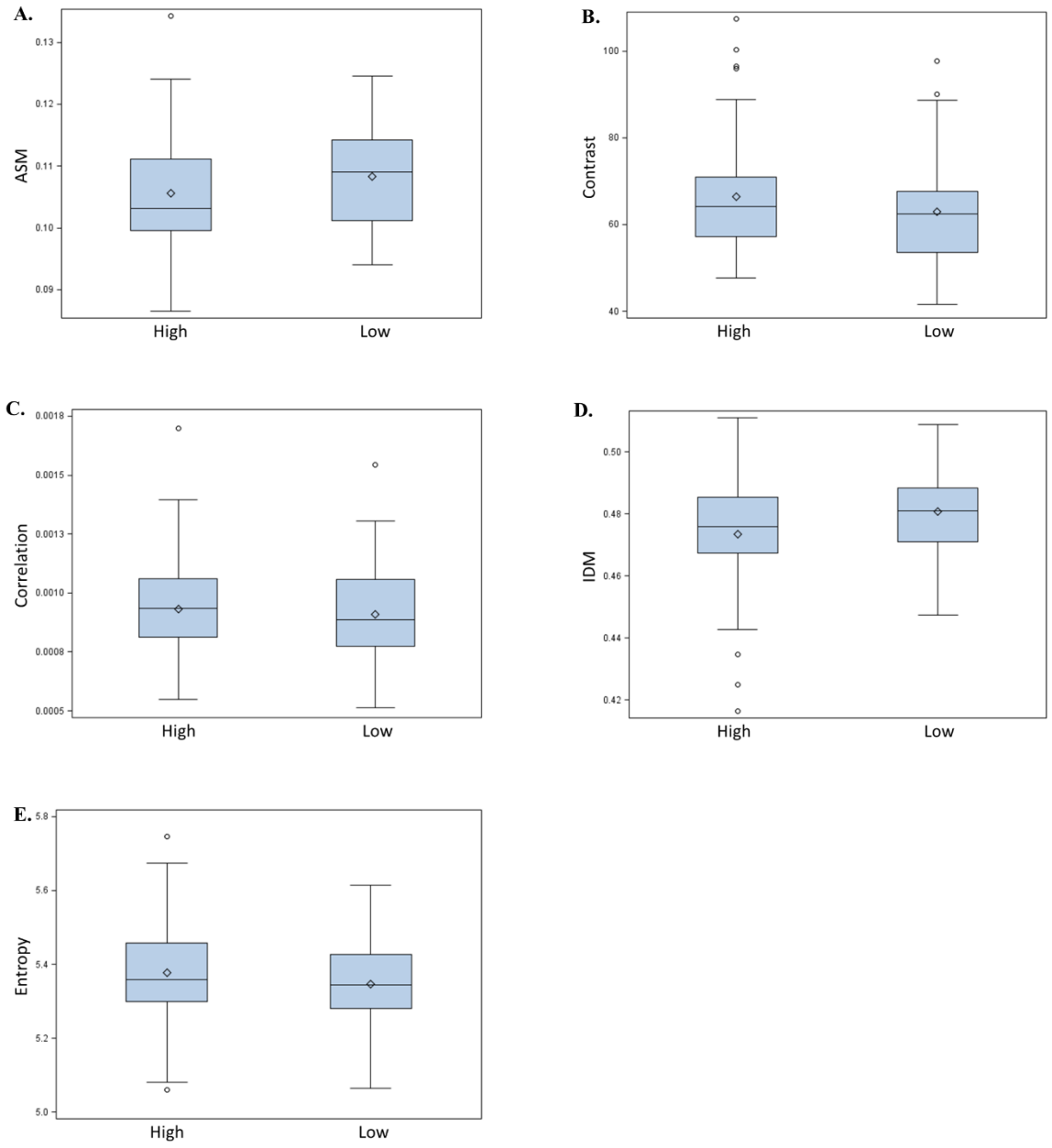
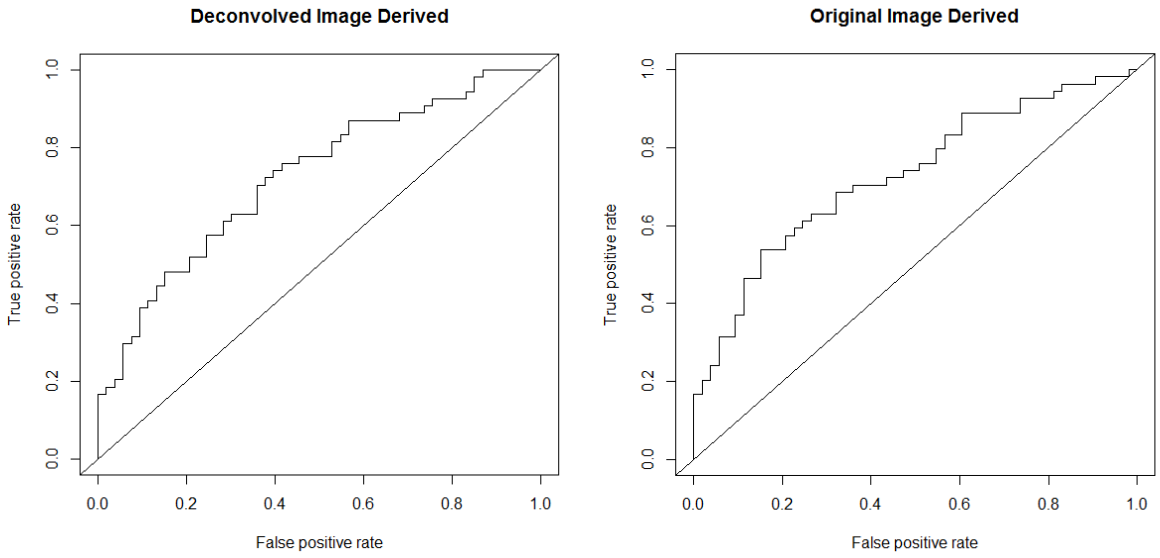


Fig 2. The box defines the interquartile range (IQR), the whiskers the last datapoints to fall within 1.5 IQR from the box, the diamond the mean and the bar defining the median. The outliers are identified as open circles located above or below the whiskers. Identified outliers were omitted from a follow-up analysis to ascertain their impact on analysis (to rule out their impact on the data as being the cause of significant difference).

Figure 3. Receiver Operator Curves (ROC) for discriminant functions classifying high and low fertility bulls from five texture parameters measured from unaltered or deconvolved nuclear fluorescent images.



Area under the curve (AUC) is 0.73 for each model, deconvolved image derived or original image derived. The diagonal line represents an AUC of 0.5, equal to that of pure chance classifying the bulls correctly. ROC's consider correct classification equally, not placing preference on one group over the other. The high degree of correct classification into the low fertility group is somewhat offset by the low degree of correct classification into the high fertility group in this graphical representation method.

Table 1.

TERMINOLOGY

P(i,j) is (i,j)th element of a co-occurrence conditional probability matrix.

N_g is the number of gray levels in the digitized images. $\mu_x, \mu_y, \sigma_x, \sigma_y$ are means and standard deviations of p_x and p_y, where x and y define rows and columns of matrices.

Textural Features

1) Angular Second Moment

$$f_1 = \sum_i \sum_j \{p(i,j)\}^2.$$

2) Contrast

$$f_2 = \sum_{n=0}^{n_{g-1}} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}_{|i-j|=n}.$$

3) Correlation

$$f_3 = \sum_{i,j} \frac{(i-\mu_x)(j-\mu_{y)}}{\sqrt{(\sigma_x \sigma_y)}} P(i,j).$$

4) Inverse Different Moment

$$f_4 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p\,(i,j).$$

5) Entropy

$$f_5 = -\sum_i \sum_j p(i,j) \log(p(i,j)).$$

Table 2.

Comparison of texture values; angle second moment (ASM), contrast (CON), correlation (COR), Inverse Different Moment (IDM) and entropy (ENT), separated by source (original or deconvolved images), and high and low fertility groups.

Texture Parameter	Original		Deconvolved	
	High	Low	High	Low
ASM	0.106 (0.001) †	0.108 (0.001)	0.106 (0.001) †	0.108 (0.001)
CON	142 (7.9) *†	171 (7.4)	66 (1.8) †	62 (1.6)
COR	5×10^{-4} (0.3×0^{-4})*†	4×10^{-4} (0.2×10^{-4})	9×10^{-4} (0.3×10^{-4})	9×10^{-4} (0.3×10^{-4})
IDM	0.522 (0.004) *†	0.509 (0.003)	0.474 (0.003) *	0.481 (0.002)
ENT	5.17 (0.03) *†	5.29 (0.03)	5.38 (0.02)	5.35 (0.01)

Means (±SEM), * indicates $p \leq 0.05$ difference between means in the two fertility groups within image type, † indicates $p \leq 0.05$ between normalized SEM of high and low fertility groups within image type. MANOVA with all five texture values indicated difference ($p < 0.05$) between fertility groups for each image type.

Table 3.

Fertility prediction outcomes for discriminant functions based on five texture values of each image type.

Actual group	Number of cases	Original Image		Deconvolved Image	
		Low	Hi	Low	High
Low	54	47 (87)	7 (13)	45 (83)	9 (17)
High	53	27 (51)	26 (49)	29 (55)	24 (45)

Percentages shown in parentheses. Both functions classify 69% into low and 31 % into high with error rates of 32% and 36 % for original and deconvolved derived functions respectively. Cross validation classified

Table 4. Canonical structure derived from discriminant analysis for texture parameters: angle second moment (ASM), contrast (CON), correlation (COR), Inverse Different Moment (IDM) and entropy (ENT),

Texture Parameter	Original	Deconvolved
ASM	0.40	0.43
CON	0.66	-0.38
COR	-0.71	-0.14
IDM	-0.68	0.59
ENT	0.68	-0.31

Squared canonical correlation values showing relationship of each texture parameter to the discriminant model. Values closest to 1 or -1 relate more strongly to the model, therefore making a larger contribution to the classification function, ultimately determining whether a sample is put into the low or high fertility group.