Received: 9 October 2021 | Revised: 16 November 2021 | Accepted: 19 November 2021

DOI: 10.1111/andr.13131

ANDROLOGY WILEY

ORIGINAL ARTICLE

From subjective to objective: A pilot study on testicular radiomics analysis as a measure of gonadal function

Bruno De Santi[1] | Giorgia Spaggiari[2] | Antonio RM Granata[2] | Marilina Romeo[2,3] | Filippo Molinari[1] | Manuela Simoni[2,3] | Daniele Santi[2,3]

[1] Biolab, Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy

[2] Unit of Endocrinology, Department of Medical Specialties, Azienda Ospedaliero-Universitaria of Modena, Ospedale Civile of Baggiovara, Modena, Italy

[3] Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy

Correspondence
Daniele Santi, Unit of Endocrinology, Department of Medical Specialties, Azienda Ospedaliero-Universitaria of Modena, Ospedale Civile of Baggiovara, Via Giardini 1355, 41126, Modena, Italy.
Email: daniele.santi@unimore.it

Abstract

Background: The connection between testicular ultrasound (US) parameters and testicular function, including both spermato- and steroidogenesis has been largely suggested, but their predictive properties are not routinely applied. Radiomics, a new engineering approach to radiological imaging, could overcome the visual limit of the sonographer.

Objectives: This study is aimed at extracting objective testicular US features, correlating with testicular function, including both spermato- and steroidogenesis, using an engineering approach, in order to overcome the operator-dependent subjectivity.

Materials and methods: Prospective observational pilot study from December 2019 to December 2020 on normozoospermic subjects and patients with semen variables alterations, excluding azoospermia. All patients underwent conventional semen analysis, pituitary-gonadal hormones assessment, and testicular US, performed by the same operator. US images were analyzed by Biolab (Turin) throughout image segmentation, image pre-processing, and texture features extraction.

Results: One hundred seventy US testicular images were collected from 85 patients (age 38.6 ± 9.1 years). A total of 44 first-order and advanced features were extracted. US inhomogeneity defined by radiomics significantly correlates with the andrologist definition, showing for the first time a mathematical quantification of a subjective US evaluation. Thirteen US texture features correlated with semen parameters, predicting sperm concentration, total sperm number, progressive motility, total motility and morphology, and with gonadotropins serum levels, but not with total testosterone serum levels. Classification analyses confirmed that US textural features predicted patients' classification according to semen parameters alterations.

Conclusions: Radiomics texture features qualitatively describe the testicular parenchyma with objective and reliable quantitative parameters, reflecting both the testicular spermatogenic capability and the action of pituitary gonadotropins. This is an innovative model in which US texture features represent a mirror of the pituitary-gonadal homeostasis in terms of reproductive function.

KEYWORDS
male infertility, radiomics, testicular function, testicular inhomogeneity, testicular ultrasound

1 | INTRODUCTION

Male gonadal dysfunctions play a causative role in the 50% of infertile couples, either alone or in combination with a female factor.[1] Despite the progressive advances made by diagnostic medicine, male infertility aetiology still remains unknown in about 30%–40% of the cases.[2]

Currently, the diagnostic framework of male infertility involves both the physical and the ultrasonography (US) testicular evaluation.[3] Testis US currently represents a topic of interest in andrological research, since several authors have tried to clarify the clinical connection between US parameters and testicular function, including both spermato- and steroidogenesis.[4,5] Recently, the European Academy of Andrology published new reference ranges for US-testicular parameters.[4,6] Among these, the testicular volume is historically the most investigated parameter, since almost 90% of it is constituted of seminiferous tubules and germ cells, thus representing an indirect measure of testicular spermatogenic capability.[7] More recently, a connection between US testicular volume and testicular steroidogenic function has been established[8,9] but the question remains whether the US testicular volume measurement is sufficient to classify testicular functions. Hitherto, a standard method to calculate US testicular volume is unavailable and reference ranges validated to predict both spermatogenic and androgen-secreting capabilities are still lacking.[4,5,8] Thus, alongside testicular volume, the role of more qualitative US parameters, such as testicular echogenicity and echostructure homogeneity/inhomogeneity, has been explored.[3,10–13] Testicular echogenicity increases during testis development, depending on seminiferous tubules' maturation and germ cell number increase.[14] Indeed, pre-pubertal testes appear more hypoechoic compared to adult ones.[14] Accordingly, testicular hypoechogenicity has been associated with reduced spermatogenesis and aberrant interstitial proliferation.[15] On the other hand, testicular US inhomogeneity, defined as the absence of a uniform structure, has been recently proposed as a marker of testicular dysfunction related to male infertility.[8] However, the definition of testis echostructure homo/inhomogeneity is operator dependent and a widely accepted quantitative measure to describe it is still not adopted.[3] The aforementioned limitations have precluded the extension of research-derived testicular US predictive properties in clinical practice. Trying to overcome these confines, Pozza et al. proposed a new scoring system aimed at predicting testicular function combining several US features in subjects referred for testicular US.[16] Although promising results has been described, the inter-operator variability still remains a weak point to objectify testis US-derived parameters, limiting all further clinically relevant applications.

Radiomics is a new engineering approach to radiological imaging, overcoming the visual limit of the radiologist by extracting numerical information from the images.[17,18] An increasing number of studies is currently available, applying radiomics in several medical field, but not in human reproduction. Bearing this in mind, a quantitative approach to assess testicular US-derived parameters is required, with the aim to reliably clarify their potential correlation with testicular function. The main aim of this pilot study is to extract objective testicular US features, correlating with testicular function, including both spermato- and steroidogenesis using a radiomics approach, in order to overcome the operator-dependent subjectivity. Engineering approaches are directly applied on testicular US images to correlate texture features with conventional semen parameters and pituitary-gonadal axis hormones, expressions of testis spermatogenetic and steroidogenic capabilities. This innovative approach to the testicular US is aimed to identify quantitatively and objective parameters able to at describing testicular US features.

2 | MATERIALS AND METHODS

A prospective observational pilot study was carried out from December 20, 2019 to December 20, 2020. All consecutive male patients attending the Andrology Unit of the Department of Medical Specialties of the Azienda Ospedaliero-Universitaria of Modena (Italy) have been considered eligible. Patients were enrolled according to the following inclusion criteria: (i) age over 18 years, (ii) attendance for couple infertility, and (iii) psycho-physical ability to sign an informed consent. Patients with diseases, genetic or not, known to alter testicular histology and/or function (i.e., hyper- or hypogonadotropic hypogonadism, previous testicular or pituitary surgery, etc.) were excluded, together with patients taking drugs interfering with hypothalamic-pituitary-testicular axis. Azoospermic patients (i.e., no sperms detected in the semen sample) were removed from the classification analysis since they could not be classified as a severe form of oligozoospermia. Accordingly, the testicular histology and the consequent US pattern could be extremely different in azoospermic patients compared to other categories. Moreover, the low number of azoospermic patients enrolled in our study group precluded the possibility to create a single category. Finally, every medical treatment for male infertility was started after the testicular US execution and seminal/hormonal basal assessments to avoid interferences on collected data.

All enrolled patients underwent andrological examination according to the clinical practice. In details, for each patient, personal and familial histories were collected, and physical examination was routinely performed. All patients underwent conventional semen analysis,[19] pituitary-gonadal hormones assessment, and testicular US. A dataset was created matching for each patient clinical, seminal, and hormonal values, while all US images were collected. All data were anonymously sent to the Biolab of the Department of Electronics and Telecommunication of Torino (Turin, Italy) for the image processing and the statistical analysis.

The present study was approved by the Ethical Committee of Modena (Protocol Number 1057/2019) and each enrolled patient provided written informed consent to participate.

2.1 | Testicular ultrasound

Testicular US was performed by a single operator using a single machine (Esaote My Lab25 Gold, Malmesbury, Wiltshire, UK) during the first evaluation, prior to knowing the hormonal and seminal patient's status. The following data from both testes were collected: Testis longitudinal and transversal sections, testicular volume, vascularization, testicular nodules detection, epididymis characteristics, varicocele, and hydrocele presence. Testicular volume was calculated using the ellipsoid formula—length (cm) × width (cm) × depth (cm) × 0.71—with measures obtained from both axial and longitudinal scans.[20] Although not definitely validated, we used this mathematical formula since its superiority in the prediction of real testicular volume was described.[21] Varicocele was graded according to Sarteschi's scale.[22] Moreover, visual inspection of parenchyma US inhomogeneity for each testis was recorded. Testicular US structure was judged by the operator following a binary classification, that is, 0 for homogeneous testis and 1 for inhomogeneous testis.

2.2 | Blood examination

After an overnight fast, morning (8.00 am) blood samples were obtained from all the patients to measure the following hormones: total testosterone, luteinizing hormone (LH), and follicle-stimulating hormone (FSH) serum levels. Serum total testosterone was evaluated by Chemiluminescent Microparticle Immunoassay (Achitect, Abbott, Dundee, UK). LH and FSH were measured by Chemiluminescent Microparticle Immunoassay (Achitect, Abbott, Longford, Ireland).

2.3 | Semen analysis

Conventional semen analysis was performed on a semen sample collected through masturbation after 2–7 days of sexual abstinence. The microscopic analysis was performed according to the world health organization criteria.[19]

2.4 | Quantitative image analysis

Radiomics applied provided the US image processing, consisting of four consecutive steps: (i) image segmentation, (ii) image pre-processing, (iii) texture features extraction, and (iv) statistical analysis. Every image processing algorithm was developed in MATLAB (Mathworks, R2020b) on a 2.21 GHz quad-core with 16 GB of RAM.

During the image segmentation (first step), the longitudinal scan was extracted from the US image (Figure 1, Panel A) and the testis was manually segmented (Figure 1, Panel B). The second step (image pre-processing) provided all strategies applied to exclude artificial image artefacts, such as measurement points, crosses, numbers, and

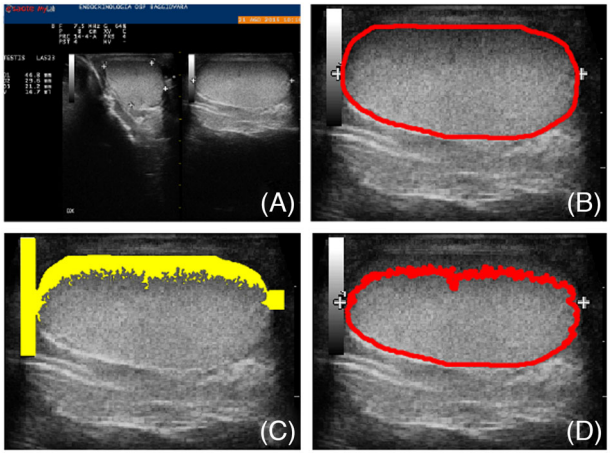

FIGURE 1 Ultrasound (US) image processing. Panel A: The testis US image obtained during a routinely performed testicular US; Panel B: The contour of the manual segmentation of the testicle in the longitudinal view; Panel C: The exclusion of calibre measurement artefacts and zones with lower echogenicity due to high acoustic impedance differences; Panel D: The final mask of the testicle for the texture analysis

zones of the parenchyma with lower echogenicity due to high acoustic impedance difference (Figure 1, Panel C). Once these artefacts were detected, they were excluded from the segmentation (Figure 1, Panel D). Finally, since US images were acquired using different gains to improve consistency in the echogenicity quantification, images were corrected using a logarithmic scale to have a uniform gain equal to 58%. Every image was amplified (in case of image with gain lower than 58 dB) or attenuated (in case of image with gain higher than 58 dB) by a factor of $20\log_{10}(|\text{gain} - 58|)/100$.

The third step provided the texture analysis, which refers to the process of extracting mathematical descriptors from an image. These descriptors are called textural features and could quantify specific properties of the image texture. In this study, a total of 44 textural features belonging to five families were extracted: (i) Four first-order features that statistically describe the distribution of pixel values including mean intensity (echogenicity), variance, skewness, and kurtosis; (ii) nine grey-level co-occurrence matrix-based (GLCM) features,[23] evaluating spatial relations between pixels with the same grey level; (iii) 13 grey-level run-length matrix-based (GLRLM) features,[24] evaluating presence of consecutive pixels with the same grey level; (iv) 13 grey-level size zone matrix-based features,[25] evaluating the presence of zones with the same grey level; (v) five neighbourhood grey tone difference matrix-based features,[26] evaluating differences between pixels with the same grey level (Table 1). Only pixel values inside the segmented area after artefacts subtraction were considered for the texture analysis.

2.5 | Statistical analysis and classification

The statistical analysis represents the fourth step of the quantitative images analysis applied with radiomics. US texture features previously described were extracted from both left and right testes, considered

separately. The statistical analysis on texture features provided first the evaluation of their reliability, evaluating the correlation between left and right testis. Second, Pearson correlation of texture features with the andrologist visual definition of inhomogeneity was performed. Third, Pearson correlation of each texture feature with conventional semen and hormone parameters was measured. In this stage, the LH on testosterone ratio was calculated to better understand the potential predictive role of US-texture images on the steroidogenic compartment. Finally, a multivariate linear regression (MLR) was applied using textural features as measurements and semen and hormone parameters as responses. R squared and *p*-values were computed. Since the analysis was performed on a large number of US textural features, the final statistical significance of Pearson correlation was adjusted using the Bonferroni correction. During these analyses, sperm motility and morphology were considered as absolute numbers and not as percentage. In order to detect any eventual pattern among the semen and hormonal parameter, a principal component analysis (PCA) was performed. The components extracted were included as responses in an MLR analysis, using US texture features as predictors.

Finally, patients were classified according to conventional semen analysis. Classification analyses were performed to measure the performance of a US texture features-based system in the detection of abnormal values of semen parameters. In details, oligozoospermia was defined when sperm concentration was lower than 15 million/ml, asthenozoospermia when progressive motility was lower than 32% and teratozoospermia when normal forms were below 4%. In order to have a robust classification system, only texture features which showed good agreement between right and left testis were considered. Further, to reduce dimensionality and complexity of the classification analysis, feature reduction was performed using PCA explaining a variance ratio of 99%. On the final dataset, classification analyses were performed by (i) multivariate linear regression analysis, and (ii) machine learning support vector machine (SVM) and artificial neural network (NN). The classification analysis was performed using a 10-fold cross-validation scheme. In details, the dataset was randomly divided into 10 folds, nine folds were randomly selected for training the classifiers and onefold for testing the classification performance. This operation was performed 10 times changing the test fold each time. This allows to evaluate the generalization ability of the classifier (i.e., how the classifier performs with unseen data). Further, to avoid any potential bias in the random selection of folds, the entire procedure was repeated 50 times.

3 | RESULTS

3.1 | Cohort characteristics

Two hundred and twenty-four US images have been collected from 112 patients (mean age 38.6 ± 9.1 years). Among these, 27 patients (54 images) were not considered for statistical analysis due to lack of hormonal or semen analyses ($n = 10$) or diagnosis of azoospermia ($n = 17$) (Figure 2). Finally, a cohort of 85 subjects (mean age was 38.4 ± 9.2 years) with seminal, hormonal, and testis US data was obtained (Figure 2). Thus, the analysis was performed on 170 US images.

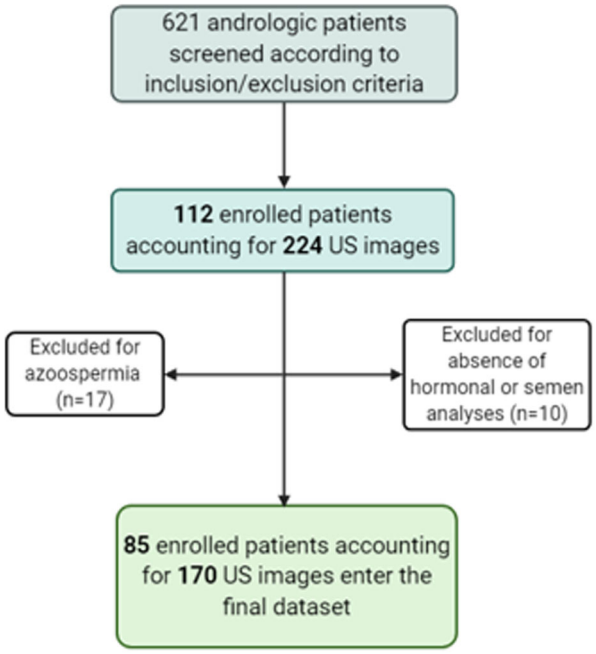

FIGURE 2 Flow chart of patients enrolled (US, ultrasonography)

The average calculated testicular volume was 15.8 ± 7.5 ml for the left and 14.8 ± 6.7 ml for the right testis. Testicular inhomogeneity was clinically described during US evaluation in 21 subjects (24.7%) for the left and in 20 subjects (23.5%) for the right testis. The presence of microlithiasis was detected in four patients (4.7%) for the left and in four (4.7%) for the right testis. No solid testicular nodules were detected in our cohort, only a testicular cyst in two patients (2.3%). The epididymal head was within reference ranges with a mean value of 8.0 ± 2.1 mm in the left and 8.6 ± 2.3 mm in the right side. Epididymal inhomogeneity was detected in 20 patients (23.5%) for the left and in 22 patients (25.9%) for the right. Hydrocele was detected in 26 patients (30.6%) and varicocele was detected in 2.3% of the dataset (two patients) in the right side and in 18.8% (16 patients) in the left side. The maximum varicocele degree detected was the 3 score in two patients.

Considering conventional semen analysis, oligozoospermia was recorded in 52.9% of the dataset (45 patients), asthenozoospermia in 56.5% (48 patients), and teratozoospermia in 69.4% (64 patients). In details, Table 1 summarizes semen analysis and hormonal value for the entire cohort analyzed.

3.2 | Texture analysis and texture features reproducibility

During the first step, the images were corrected for consistence of echogenicity. This step was fundamental to adjust the different gains used to acquire the snapshot. Then, longitudinal

TABLE 1 Patient's characteristics considering both semen and hormonal examinations

Parameter	Reference range	Value
Semen analysis		
Semen volume (ml)	>1.5	2.5 (1.9)
Semen pH	>7.2	8.0 ± 0.3
Sperm concentration (million/ml)	>15	17.3 (18.0)
Total sperm number (million)	>39	41.3 (38.1)
Progressive sperm motility (%)	>32	20.0 (40.0)
Total sperm motility (%)	>40	30.0 (43.0)
Normal forms (%)	>4	1.0 (4.0)
Hormonal assessment		
Testosterone (ng/ml)	2.2–7.8	5.1 ± 2.4
LH (IU/L)	1–9	3.5 (2.5)
FSH (IU/L)	1–12	4.6 (4.5)
Clinical characteristics		
Unilateral varicocele, *n*(%)	–	14 (16.5%)
Bilateral varicocele, *n*(%)	–	2 (2.3%)
History of cryptorchidism, *n*(%)	–	1 (1.2%)

Data are expressed as mean ± standard deviation or median (interquartile range).

scans were analyzed by the engineering approach. During the pre-processing phase, each image underwent segmentation and artefacts removal (Figure 1). The second phase provided the texture features extraction. As reported in the methods section, both first-order and advanced features were extracted for a total of 44 variables.

Since each testis was considered separately, the reproducibility of texture features between left and right testis in the same patient was first evaluated. A threshold of 0.5 on the left/right correlation was used to define an acceptable reproducibility. With this approach, 35 on 44 US texture features showed good agreement between right and left testis. This result suggests that the texture features extraction is a good model to describe testicular US characteristics in the same patient, with a good reproducibility between the two testes.

3.3 | Correlation with visual defined inhomogeneity

The US inhomogeneity defined by the engineering approach was statistically significant correlated with the inhomogeneity defined by the andrologist. In particular, after Bonferroni correction, 19 texture features were identified as a predictor of the clinical definition of inhomogeneity (Table 2). Further, multiple linear regression analysis highlighted that US texture features significantly predict visually defined inhomogeneity (*R*-squared 0.379, $p < 0.001$).

3.4 | Correlation with semen parameters

Thirteen US texture features were significantly correlated with semen parameters (Table 3). In particular, 12 US texture features correlated with total sperm number, 12 texture features with progressive motility, and 12 texture features with total motility (Table 3). Interestingly, no US texture features correlated with sperm morphology (Table S1).

Multiple linear regression analyses highlighted that US texture features significantly predict sperm concentration (*R*-squared 0.336, $p = 0.041$), total sperm number (*R*-squared 0.380, $p = 0.009$), progressive motility (*R*-squared 0.343, $p = 0.046$), total motility (*R*-squared 0.358, $p = 0.025$), and morphology (*R*-squared 0.397, $p = 0.005$). Surprisingly, despite the univariate correlation of texture features with morphology detected a low degree of association, multivariate analysis showed good prediction of morphology using US texture features. These results confirm that the US texture features could have a role in the spermatogenic pattern prediction.

The analysis was repeated considering only those US texture features that correlated with the visual US inhomogeneity. With this adjustment, the US texture features significantly predict only sperm concentration (*R*-squared 0.195, $p = 0.017$), but not total sperm count (*R*-squared 0.165, $p = 0.073$), progressive motility (*R*-squared 0.147, $p = 0.155$), total motility (*R*-squared 0.149, $p = 0.141$), and morphology (*R*-squared 0.117, $p = 0.420$). This result suggests that the engineering evaluation of the US inhomogeneity provides more detailed information than the clinician definition.

3.5 | Texture features analysis and correlation with hormonal parameters

Considering pituitary-gonadal hormones, US texture features did not significantly correlate with total testosterone serum levels. On the contrary, US parameters significantly correlated with gonadotropins serum levels. In particular, 28 US texture features correlated with LH and 30 with FSH (Table 4). Moreover, interestingly, 19 US texture features correlated with LH/testosterone ratio (Table 4).

Multivariate analyses combined US-texture features and semen parameters (sperm concentration, total sperm number, progressive and total sperm motilities, and typical forms) and hormonal variables (testosterone, LH, and FSH). These analyses confirmed that US texture features significantly predicted LH (*R*-squared 0.713, $p < 0.001$) and FSH (*R*-squared 0.656, $p < 0.001$), while did not predict testosterone serum levels (*R*-squared 0.267, $p = 0.131$). Using only those US texture features correlated with the clinician

TABLE 2 Correlation analysis between ultrasound (US) texture features and visual inhomogeneity as defined by the andrologist

Features	Pearson Correlation	*p*-Value	Beta coefficient
Mean	**−0.34**	**<0.001**	−0.03
Skewness	**0.31**	**<0.001**	0.03
GLCM_SumAverage	**−0.30**	**<0.001**	−2.47
GLCM_AutoCorrelation	**−0.29**	**<0.001**	11.50
LGRE	**0.32**	**<0.001**	0.42
HGRE	**−0.28**	**<0.001**	25.80
SRLGE	**0.32**	**<0.001**	0.23
SRHGE	**−0.28**	**<0.001**	−35.20
LRLGE	**0.26**	**<0.001**	1.09
LRHGE	**−0.30**	**<0.001**	0.59
GLVR	**−0.31**	**<0.001**	0.10
LGZE	**0.31**	**<0.001**	−0.39
HGZE	**−0.28**	**<0.001**	1.46
SZLGE	**0.30**	**<0.001**	−1.14
SZHGE	**−0.28**	**<0.001**	−0.13
LZHGE	**−0.29**	**<0.001**	−2.19
GLVZ	**−0.25**	**<0.001**	−0.01
ZSV	**0.25**	**<0.001**	−0.12
Strength	**0.26**	**<0.001**	0.07

Note: Bold values represent parameters significantly correlated with visual US inhomogeneity. The last column shows the beta coefficients of the multivariate analysis.

Abbreviations: GLCM, grey-level co-occurrence matrix; GLN, grey-level non-uniformity; GLV, grey-level-variability; GLVR, grey-level variability of runs; GLVZ, grey-level variability of zones; GLVZ, grey-level variability of zones; HGRE, high grey-level run emphasis; HGZE, high grey-level zone emphasis; LGRE, low grey-level run emphasis; LGZE, low grey-level zone emphasis; LRE, long run emphasis; LRHGE, long run high grey-level emphasis ; LRLGE, long run low grey-level emphasis; LZHGE, large zone high grey-level emphasis; LZE, length size emphasis; LZLGE, large zone low grey level emphasis; RP, run percentage; RLN, run length non-uniformity; RL, run length velocity; SRHGE, short run high grey-level emphasis; SRE, short run emphasis; SRLGE, short run low grey-level emphasis; SZHGE, small zone high grey-level emphasis; SZE, small zone emphasis; SZLGE, small zone low grey-level emphasis; ZP, zone percentage; ZSN, zone size non-uniformity; ZSV, zone size variability.

visual US inhomogeneity, the significant prediction was confirmed for LH (*R*-squared 0.326, $p < 0.001$) and FSH (*R*-squared 0.357, $p < 0.001$), and not for testosterone serum levels (*R*-squared 0.127, $p = 0.286$).

3.6 | Principal component analysis

The PCA was performed considering semen parameters and hormonal asset. The PCA detected three significant components (Figure 3). The first component (PC1) enclosed all semen parameters (variance 55%), the second (PC2) gonadotropins (variance 23%), and the third (PC3) testosterone serum levels (variance 12%) (Figure 3). The three components were then used in multivariate linear regression analysis with US texture features. Interestingly, US texture features extracted significantly predicted PC1 (*R*-squared 0.393, $p = 0.017$) and PC2 (*R*-squared 0.452, $p < 0.001$), but not PC3 (*R*-squared 0.298, $p = 0.350$). These results confirmed that the engineering extraction of US features could predict semen parameters and gonadotropins levels, but not testosterone secretion.

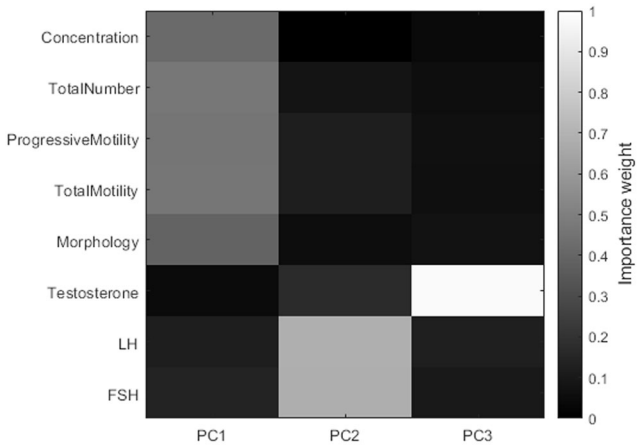

FIGURE 3 Principal component analysis (PCA) considering semen parameters, gonadotropins, and testosterone serum levels. Importance weights of features into principal components are encoded in a colourmap from black (null weight) to white (weight equal to 1) (FSH = follicle-stimulating hormone; LH = luteinizing hormone; PC = principal component)

TABLE 3 Correlation analysis between ultrasound texture features and semen parameters obtained by conventional semen analysis

	Semen parameters											
	Concentration			Total number			Progressive motility			Total motility		
Features	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta
Mean	0.2	**0.009**	14.5	0.2	**0.020**	24.3	0.1	0.119	2261.3	0.1	0.075	1918.4
Variance	−0.1	0.251	1.1	−0.1	0.412	37.5	−0.1	0.482	−223.9	0.0	0.537	376.3
Skewness	−0.1	0.041	−17,470.0	−0.1	0.059	−25.1	-0.1	0.234	−97916.0	−0.1	0.157	−185297.0
Kurtosis	0.1	0.213	27,861.0	0.0	0.487	−9.3	0.0	0.925	−144,060.0	0.0	0.818	−31,256.0
GLCM_Energy	−0.1	0.102	−3032.0	-0.1	0.111	144.3	−0.1	0.116	−15,827.0	−0.1	0.123	−1,821,791.0
GLCM_Contrast	0.3	<**0.001**	−10,387.0	0.3	<**0.001**	−347.3	0.3	<**0.001**	−821,212.0	0.3	**0.000**	−192,075.0
GLCM_Entropy	0.1	0.082	−10,964.0	0.1	0.094	284.8	0.1	0.097	256,873.0	0.1	0.101	078,507.0
GLCM_Homogeneity	-0.3	<**0.001**	449,981.0	−0.3	<**0.001**	174.5	−0.2	**0.001**	5,328,277.0	−0.2	**0.001**	7,023,606.0
GLCM_Correlation	−0.2	**0.002**	136,775.0	-0.2	**0.002**	−21.8	−0.2	**0.002**	−612,505.0	−0.2	**0.002**	-606,790.0
GLCM_SumAverage	0.1	0.111	515,620.0	0.1	0.254	334.5	0.0	0.510	2,755,099.0	0.1	0.394	1,843,415.0
GLCM_Variance	0.0	0.650	266,897.0	0.0	0.706	−4.4	0.0	0.716	−401,372.0	0.0	0.751	−363,877.0
GLCM_Dissimilarity	0.3	<**0.001**	79,253.0	0.3	<**0.001**	−10.0	0.3	<**0.001**	−1,917,661.0	0.2	<**0.001**	−74,734.0
GLCM_AutoCorrelation	0.1	0.163	−150,463.0	0.1	0.341	−323.0	0.0	0.621	986,885.0	0.0	0.496	3,878,548.0
SRE	0.3	<**0.001**	-48,162.0	0.3	<**0.001**	-312.0	0.3	<**0.001**	−1,940,323.0	0.3	<**0.001**	−1,590,729.0
LRE	−0.2	**0.001**	−309,302.0	−0.2	**0.002**	−1218.4	−0.2	**0.005**	−2,476,272.0	−0.2	**0.005**	−2,823,157.0
GLN	0.0	0.518	−108,844.0	0.0	0.574	−421.2	0.0	0.663	−810,569.0	0.0	0.689	−1,550,265.0
RLN	0.3	<**0.001**	58,811.0	0.3	<**0.001**	214.7	0.3	<**0.001**	103,817.0	0.3	<**0.001**	1,037,187.0
RP	0.2	<**0.001**	−211,373.0	0.3	<**0.001**	−799.2	0.2	**0.001**	2,022,165.0	0.2	**0.001**	3,106,255.0
LGRE	−0.1	0.060	109,036.0	−0.1	0.118	181.0	−0.1	0.272	1,012,395.0	−0.1	0.210	−1,821,785.0
HGRE	0.1	0.126	851,893.0	0.1	0.222	205.6	0.1	0.470	−20,838.4	0.1	0.355	−51,203.9
SRLGE	−0.1	0.180	−61.8	−0.1	0.320	91.4	0.0	0.542	26,191.0	-0.1	0.468	−6045.6
SRHGE	0.2	**0.008**	−54.5	0.2	**0.013**	−46.1	0.1	**0.039**	326.2	0.2	**0.025**	−2607.8
LRLGE	−0.2	**0.025**	−8.0	−0.1	0.064	−55.2	−0.1	0.169	−44,982.0	−0.1	0.140	−50,251.9
LRHGE	−0.1	0.090	−94.2	−0.1	0.057	−142.7	−0.1	0.047	−4526.6	−0.1	0.073	−2822.7
GLVR	0.3	<**0.001**	43.3	0.3	<**0.001**	205.6	0.3	<**0.001**	16,769.6	0.3	<**0.001**	16,630.7
RLV	0.2	**0.001**	−4.0	0.3	<**0.001**	−33.5	0.2	**0.001**	−4541.7	0.2	**0.001**	−4143.6

(Continues)

TABLE 3 (Continued)

	Semen parameters											
	Concentration			Total number			Progressive motility			Total motility		
Features	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta
SZE	0.3	<**0.001**	−59.5	0.3	<**0.001**	−274.2	0.3	<**0.001**	−18,236.3	0.3	<**0.001**	−23,257.3
LZE	−0.2	**0.003**	74.4	−0.2	**0.007**	336.2	−0.2	0.017	25,191.3	−0.2	0.018	28,223.0
GLN	0.0	0.524	64.3	0.0	0.538	430.1	0.0	0.627	18,530.1	0.0	0.650	31,708.9
ZSN	0.3	<**0.001**	80.0	0.3	<**0.001**	363.7	0.3	<**0.001**	23,812.5	0.3	<**0.001**	29,555.9
ZP	0.3	<**0.001**	27.7	0.3	<**0.001**	141.1	0.3	<**0.001**	10,028.8	0.3	<**0.001**	10,206.7
LGZE	−0.1	0.062	15.2	−0.1	0.142	−624.7	−0.1	0.322	−48,764.2	−0.1	0.250	42,335.9
HGZE	0.1	0.147	40.7	0.1	0.229	−340.2	0.0	0.500	−22,147.0	0.1	0.356	−19,829.7
SZLGE	−0.1	0.189	−40.4	−0.1	0.334	257.8	0.0	0.572	16,419.2	0.0	0.486	105.1
SZHGE	0.2	**0.019**	−1.6	0.1	0.049	33.7	0.1	0.112	3491.3	0.1	0.080	3784.3
LZLGE	−0.1	0.198	−17.1	−0.1	0.255	−204.5	−0.1	0.351	19,764.8	−0.1	0.333	5579.8
LZHGE	−0.1	0.097	97.4	−0.1	0.094	164.7	−0.1	0.051	3639.9	−0.1	0.133	5497.4
GLVZ	0.3	0.000	4.5	0.3	**0.000**	21.0	0.2	**0.001**	−58.6	0.2	**0.001**	1066.0
ZSV	−0.2	0.004	13.7	0.2	0.007	25.8	0.2	0.009	793.0	0.2	0.009	580.4
Coarseness	−0.2	0.007	0.6	-0.1	0.044	−19.8	−0.1	0.112	−650.8	−0.1	0.101	−1136.1
Contrast	0.1	0.052	−4.4	0.2	0.018	4.1	0.2	0.014	4286.9	0.2	0.015	4968.1
Busyness	0.2	0.033	37.9	0.1	0.054	129.0	0.1	0.047	4993.6	0.1	0.063	5926.8
Complexity	0.2	0.011	37.5	0.2	0.029	81.2	0.2	0.030	7951.8	0.2	0.032	9755.7
Strength	−0.2	0.015	−5.7	−0.1	0.062	23.1	−0.1	0.126	1747.4	−0.1	0.122	1905.2

Note: Bold values represent significant correlation values.

Abbreviations: GLCM, grey-level co-occurrence matrix; GLN, grey-level non-uniformity; GLV, grey-level-variability; GLVR, grey-level variability of runs; GLVZ, grey-level variability of zones ; GLVZ, grey-level variability of zones; HGRE, high grey-level run emphasis; HGZE, high grey-level zone emphasis; LGRE, low grey-level run emphasis; LGZE, low grey-level zone emphasis; LRE, long run emphasis; LRHGE, long run high grey-level emphasis ; LRLGE, long run low grey-level emphasis; LZHGE, large zone high grey-level emphasis; LZE, length size emphasis; LZLGE, large zone low grey level emphasis; RP, run percentage; RLN, run length non-uniformity; RL, run length velocity; SRHGE, short run high grey-level emphasis; SRE, short run emphasis; SRLGE, short run low grey-level emphasis; SZHGE, small zone high grey-level emphasis; SZE, small zone emphasis; SZLGE, small zone low grey-level emphasis; ZP, zone percentage; ZSN, zone size non—uniformity; ZSV, zone size variability.

TABLE 4 Correlation analysis between ultrasound texture features and hormones

	Pituitary-testicular hormonal profile								
	Testosterone			LH			FSH		
Features	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta
Mean	0.0	0.679	0.6	−0.4	<**0.001**	−0.7	−0.5	<**0.001**	−1.5
Variance	−0.1	0.217	−1.6	0.1	0.065	0.6	0.1	0.348	4.3
Skewness	0.0	0.892	0.1	0.4	<**0.001**	0.2	0.4	<**0.001**	1.6
Kurtosis	0.0	0.798	0.2	−0.1	0.105	0.2	−0.1	0.237	1.1
GLCM_Energy	−0.1	0.432	1.7	−0.1	0.277	0.2	0.0	0.492	3.2
GLCM_Contrast	0.1	0.125	559.8	−0.2	**0.009**	5.2	−0.2	**0.014**	−11.8
GLCM_Entropy	0.1	0.232	2.9	−0.1	0.488	0.7	−0.1	0.260	7.9
GLCM_Homogeneity	−0.1	0.086	−1762.5	0.3	<**0.001**	−2.2	0.2	**0.001**	25.1
GLCM_Correlation	−0.1	0.266	0.1	0.3	<**0.001**	1.0	0.2	**0.002**	3.4
GLCM_SumAverage	0.0	0.722	1.7	−0.3	<**0.001**	5.4	−0.3	<**0.001**	2.6
GLCM_Variance	0.1	0.218	1.3	0.1	0.114	−0.6	0.1	0.339	−3.4
GLCM_Dissimilarity	0.1	0.099	−2322.8	−0.2	0.005	−6.9	−0.2	0.011	32.5
GLCM_AutoCorrelation	0.0	0.609	−6.3	−0.2	**0.001**	−1.6	−0.3	<**0.001**	−23.2
SRE	0.1	0.041	−0.8	−0.5	<**0.001**	0.5	−0.4	<**0.001**	3.9
LRE	−0.1	0.067	−2.7	0.4	<**0.001**	19.7	0.3	<**0.001**	40.0
GLN	0.0	0.806	−0.5	−0.2	**0.011**	−3.9	−0.1	0.063	−10.2
RLN	0.1	0.068	2.2	−0.4	<**0.001**	−0.6	−0.3	<**0.001**	−3.8
RP	0.1	0.093	−7.9	−0.3	<**0.001**	20.9	−0.3	<**0.001**	49.9
LGRE	0.0	0.777	3.7	0.4	<**0.001**	0.9	0.4	<**0.001**	−4.1
HGRE	0.0	0.552	0.8	−0.2	**0.002**	30.4	−0.3	<**0.001**	66.1
SRLGE	0.1	0.187	−1.0	0.4	<**0.001**	8.2	0.4	<**0.001**	−22.5
SRHGE	0.0	0.622	−0.7	−0.3	<**0.001**	-0.8	−0.3	<**0.001**	−6.0
LRLGE	−0.1	0.456	−0.1	0.5	<**0.001**	118.9	0.4	<**0.001**	141.2
LRHGE	−0.1	0.174	0.3	−0.2	0.003	1.0	−0.2	**0.001**	−16.8
GLVR	−0.1	0.404	−0.8	−0.3	<**0.001**	−0.6	−0.4	<**0.001**	−0.9
RLV	0.1	0.155	1.0	−0.3	<**0.001**	0.3	−0.3	<**0.001**	−0.4
SZE	0.1	0.067	−7.8	−0.4	<**0.001**	−1.3	−0.4	<**0.001**	−6.5
LZE	−0.1	0.067	−3.9	0.4	<**0.001**	0.4	0.3	<**0.001**	−18.5
GLN	0.0	0.533	−0.9	−0.2	0.017	4.2	−0.1	0.058	16.3
ZSN	0.1	0.075	8.6	−0.4	<**0.001**	2.1	−0.4	<**0.001**	6.9
ZP	0.1	0.090	−0.1	−0.3	<**0.001**	−1.0	−0.3	<**0.001**	−18.7
LGZE	0.0	0.549	−3.2	0.4	<**0.001**	−61.6	0.4	<**0.001**	−153.8
HGZE	−0.1	0.382	0.1	−0.2	**0.002**	−31.8	−0.3	<**0.001**	−22.5
SZLGE	0.1	0.205	0.6	0.3	<**0.001**	−4.2	0.3	<**0.001**	73.1
SZHGE	−0.1	0.477	3.9	−0.3	<**0.001**	−2.1	−0.3	<**0.001**	−0.8
LZLGE	−0.1	0.161	0.0	0.5	<**0.001**	−89.5	0.5	<**0.001**	−18.2
LZHGE	−0.1	0.054	−0.2	−0.1	0.039	−0.2	−0.2	0.008	−0.8
GLVZ	0.0	0.495	0.0	−0.3	<**0.001**	0.7	−0.3	<**0.001**	−0.7
ZSV	−0.1	0.139	1.1	0.6	<**0.001**	−0.3	0.5	<**0.001**	0.3
Coarseness	0.0	0.625	0.5	0.3	<**0.001**	−0.4	0.3	<**0.001**	−0.5

(Continues)

TABLE 4 (Continued)

	Pituitary-testicular hormonal profile								
	Testosterone			LH			FSH		
Features	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta	Correlation	*p*-Value	Beta
Contrast	0.1	0.178	−0.1	−0.1	0.110	−0.6	−0.2	**0.032**	−4.8
Busyness	0.0	0.826	−0.2	−0.2	**0.027**	−0.6	−0.1	0.051	−1.7
Complexity	0.2	0.028	0.2	−0.1	0.206	−1.9	−0.1	0.177	−3.1
Strength	0.1	0.117	−0.7	0.4	<**0.001**	1.4	0.3	<**0.001**	1.9

Bold values represent significant value.

Abbreviations: GLCM, grey-level co-occurrence matrix; GLN, grey-level non-uniformity; GLV, grey-level-variability; GLVR, grey-level variability of runs; GLVZ, grey-level variability of zones; GLVZ, grey-level variability of zones; HGRE, high grey-level run emphasis; HGZE, high grey-level zone emphasis; LGRE, low grey-level run emphasis; LGZE, low grey-level zone emphasis; LRE, long run emphasis; LRHGE, long run high grey-level emphasis; LRLGE, long run low grey-level emphasis; LZHGE, large zone high grey-level emphasis; LZE, length size emphasis; LZLGE, large zone low grey level emphasis; RP, run percentage; RLN, run length non-uniformity; RL, run length velocity; SRHGE, short run high grey-level emphasis ; SRE, short run emphasis; SRLGE, short run low grey-level emphasis; SZHGE, small zone high grey-level emphasis; SZE, small zone emphasis; SZLGE, small zone low grey-level emphasis; ZP, zone percentage; ZSN, zone size non−uniformity; ZSV, zone size variability.

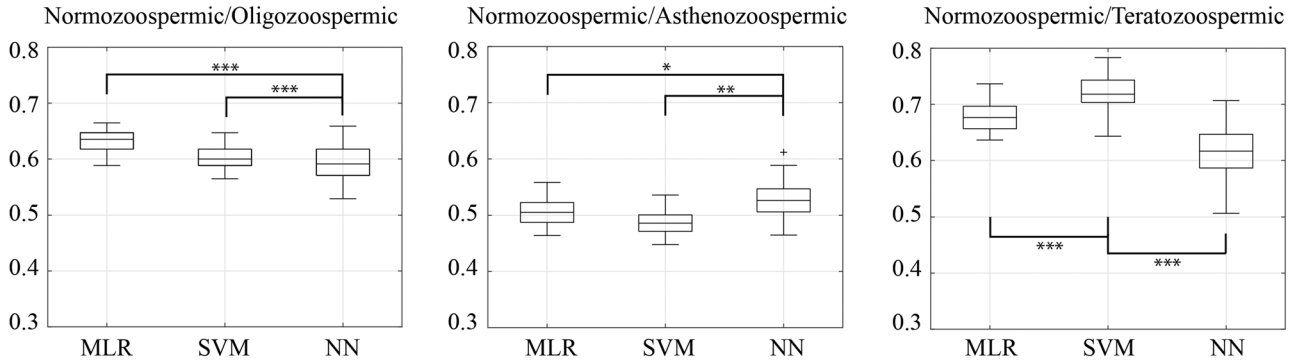

FIGURE 4 Classification performance for differentiation between normozoospermic and oligozoospermic, normozoospermic and asthenozoospermic, normozoospermic, and teratozoospermic patients using multivariate linear regression (MLR), support vector machine (SVM), and neural network (NN) classifier. Row represents area under the curve (AUC) values in the form of diagram bars with mean and standard deviation values. $^{*}p < 0.05$, $^{**}p < 0.001$, $^{***}p < 0.0001$ (ANOVA 1-way with Bonferroni correction)

3.7 | Classification analyses

During classification analysis, the first step provided the reduction of dataset complexity using PCA. Thus, the 35 reproducible features initially extracted were reduced to 15 principal components. Considering the classification in normo- and oligozoospermic patients, MLR and SVM classifiers better worked, with Area under the curve 0.62 ± 0.02 and 0.61 ± 0.02, respectively (Figure 4). Performances were lower when patients were classified according to sperm motility, with AUC values around 0.5 for every classifier (Figure 4). Finally, considering sperm morphology classification, SVM better worked, with AUC 0.73 ± 0.03 (Figure 4).

Finally, classification analyses were performed adjusting data using the US-calculated testicular volume. The classification performance significantly increased for normo-/oligozoospermia using the MLR classifier (MLR: from 0.62 ± 0.02 to 0.65 ± 0.02) but not for SVM and NN and for normo-/teratozoospermia differentiation using every classifier (MLR: from 0.64 ± 0.03 to 0.72 ± 0.02, SVM: from 0.73 ± 0.03 to 0.81 ± 0.02, NN: from 0.62 ± 0.04 to 0.70 ± 0.05). This result suggests that US-derived testicular volume could have a role in predicting only the sperm number.

4 | DISCUSSION

This is the first radiomics application on testicular US images with the aim to objectify an instrumental examination burdened by high variability and low accuracy due to operator dependence. Here, we can extract the texture features from US images, qualitatively describing the testicular parenchyma with objective and reliable quantitative parameters. On the one hand, these echostructure parameters can objectively describe the testicular parenchyma; on the other hand, they reflect testicular spermatogenic capability and are in turn influenced by pituitary gonadotropins. Thus, we identified an innovative model formed by specific US texture features representing a mirror of the pituitary-gonadal homeostasis in terms of reproductive function.

The US testis evaluation during the diagnostic work-up of male infertility provides a great deal of information to the clinician. However, the vast majority of the obtained parameters gives only a qualitative description of the male gonad, burdened by both intra- and inter-operator variability, and in general by the operator subjectivity. However, several attempts to predict the hormonal and seminal status from the US testicular echotexture have been performed so far. A normally functioning testis shows a homogeneous echotexture, characterized by homogenously distributed medium-level echoes,[3] whereas an impaired function is generally associated to US inhomogeneity, defined as the presence of areas of altered echogenicity, mostly hypoechoic.[15] Indeed, the detection of a reduced US echogenicity seemed related to either reduced spermatogenesis or aberrant interstitial proliferation, justifying an impairment in physiological reproductive functions.[27] Accordingly, in experimental model, testicular echostructure alterations have been associated to specific histological findings, such as fibrosis, tubular sclerosis, and spermatogenic arrest.[28,29] However, the subjectivity of the US evaluation still remains, limiting the quantification of the relationship between US alterations and testicular function. Lenz et al. tried to objectify the qualitative-expressed testicular homo/inhomogeneity, developing a 5-point score in which a higher score reflected irregular echotexture.[22,30,31] Recently, a semi-quantitative scoring system was proposed, matching all parameters obtained during testicular US in a logistic regression analysis, predicting both impaired spermatogenesis and hypogonadism better than Lenz's score.[16] However, this interesting algorithm did not overcome the operator subjectivity at defining the testicular US echostructure. In this context, innovative approaches, such as radiomics, could be extremely useful to combine qualitative and quantitative information coming from testicular US examination. Our study is the natural continuation of this line of research, applying more complex engineering analyses directly on US images, trying to extract accurate and informative parameters able to describe the testicular echotexture. In our study, a single US image provides more than 40 different echostructure parameters, among which 19 correlate with the visual US inhomogeneity defined by the clinician during the examination. The multivariate analysis showed a significant prediction of the visual US inhomogeneity by the texture features. This result is extremely relevant and innovative. Indeed, this is the first mathematical and engineering quantification of a subjective US evaluation. The radiomics is able to reproduce and overlap the operator's judgment.

Even applying more complex mathematical algorithms, the testis US echostructure is confirmed able to predict the spermatogenic capability, measured through conventional semen analysis. Indeed, 12 US texture features correlate and predict both sperm number and motility. Intriguingly, none of the US features result able to predict sperm morphology. Thus, the detection of testis US homogeneity/inhomogeneity results in an indirect marker of spermatogenetic status, at least in terms of sperm number and motility. However, when all semen parameters have been considered together, multivariate and classification analyses confirmed the predictive ability of US texture to classify patients according to semen abnormalities, including sperm morphology. The radiomics ability to describe tissue structure and to predict tissue function has been recently confirmed in different contexts, such as thyroid,[32,33] parathyroid,[34] breast,[35,36] kidney,[37] and prostate.[38] Here, for the first time, we applied this engineering technology to the testis, confirming that US texture features could be useful in clinical practice to monitor testicular function during infertility work-up. Future studies are required to increase the accuracy of the mathematical model applied, reaching the complete prediction ability of US analysis. In particular, future studies should increase the size of the dataset applied to machine learning methods, increase the number of US texture features extracted, and test deep learning algorithms for automatic and complex feature extraction.

Intriguingly, our study highlights a specific set of US texture features which correlates with all semen and hormonal parameters, showing high reproducibility when evaluated in left and right testis. However, no first-order features (i.e., mean and variance) are present among these features. Thus, advanced texture features better correlate with the clinical picture of fertility, at least in our application. One of these is homogeneity derived from the GLCM matrix which describes the presence of neighbouring pixels with similar grey levels, measuring the grey levels dispersion of neighbouring pixels. The advanced US texture features here detected belong to the GLRLM family, evaluating the presence of runs of consecutive pixels having the same grey level. In particular, short run emphasis (SRE) and long run emphasis (LRE) count how many short or long runs are there, while run percentage (RP) describes how many runs are and grey-level variability of runs measures the variability of grey level among different runs. Finally, the other texture features comprehensively evaluate the presence of zones with the same grey level. Thus, every US texture feature descriptor is a mathematical descriptor of inhomogeneity, granularity and coarseness of an image, using different algorithms. The wide number of parameters detected in correlation with the visual clinician judgement of inhomogeneity confirms the complexity that lies behind the US operator's experience and subjectivity. Here, these parameters prove effective in quantifying the extreme subjectivity inherent in the testicular US.

Considering the androgen-secreting compartment, our US radiomics approach does not predict testosterone serum levels. Thus, a comprehensive testis US-derived model able to predict both spermatogenetic and steroidogenic functions seem to be unreachable. However, total testosterone serum levels are the final result of both gonadal and adrenal production,[39] balanced by the liver production of albumin and sex hormone binding globulin. Thus, not so unexpectedly, the testis US features analysis alone is unable to comprehensively predict the final testosterone serum levels, notwithstanding the AI application. However, testicular echostructure features correlate with gonadotropins serum levels. These glycoprotein pituitary hormones have an exquisitely testicular bidirectional tropism. In particular, the testicular target of LH is the Leydig cell, deputed to testosterone production.[40] However, androgens production is not a process unique to testicular tissue, since it occurs also in adrenal glands under different regulatory mechanisms, that is, in response to adrenocorticotropic hormone.[41] With this in mind, the correlation between testicular US features and LH serum levels could be interpreted as a more accurate prediction of the Leydig cells component, rather than simply

testosterone serum levels. Accordingly, when LH/testosterone ratio was considered, the correlation was confirmed, considering the strict relationship between US features and steroidogenic compartment. On the other hand, the interconnection between FSH and testicular sonographic pattern reinforces the US capability to be informative about the testicular spermatogenetic function. Comprehensively, we could speculate that US radiomics results are able to predict both the seminiferous and the interstitial testicular components.

Considering the innovative approach to testicular US images, our approach shows several limits. First, we enrolled patients evaluated for couple infertility. Although among these patients we detected subjects without andrological problems and with normal semen and hormonal parameters, the use of strict inclusion criteria limits the exportability of the result. Second, this study applied radiomics to testicular US, but it is not able to detect an immediate application of this technique in clinical practice. Indeed, this is the first step that could help researchers to convert digital images to mineable data that could be used in future studies to build a clinical-engineering tool. In this setting, overall, a true control group including healthy fertile men is lacking. Thus, this study should be considered as a pilot study for future potential applications.

In conclusion, this is the first attempt to match radiomics and gonadal male function. This study could not be directly translated into an immediate clinical message. Indeed, larger case series are needed to confirm and to ameliorate the model's performance. This study shows the potentiality to impact the management of the andrological patient. The real advantage of this study is the conversion of subjective parameters, obtained by testicular US examination, to objective, numerical data, that could be evaluated using a statistical approach.

ACKNOWLEDGMENTS

None.

FUNDING INFORMATION

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Daniele Santi conceived the idea. Giorgia Spaggiari, Antonio RM Granata and Daniele Santi collected images, Bruno D Santi and Filippo Molinari performed engineering analyses, Bruno D Santi, Giorgia Spaggiari and Daniele Santi wrote the manuscript, Manuela Simoni revised the manuscript.

ORCID

Bruno De Santi https://orcid.org/0000-0003-3790-7645
Manuela Simoni https://orcid.org/0000-0002-2133-4304

REFERENCES

1. de Kretser DM. Male infertility. *Lancet London Engl*. 1997;349(9054):787-790.
2. Vockel M, Riera-Escamilla A, Tüttelmann F, Krausz C. The X chromosome and male infertility. *Hum Genet*. 2021;140:203-215.
3. Lotti F, Maggi M. Ultrasound of the male genital tract in relation to male reproductive health. *Hum Reprod Update*. 2015;21(1):56-83.
4. Lotti F, Frizza F, Balercia G, et al. The European Academy of Andrology (EAA) ultrasound study on healthy, fertile men: clinical, seminal and biochemical characteristics. *Andrology*. 2020;8(5):1005-1020.
5. Lotti F, Frizza F, Balercia G, et al. The European Academy of Andrology (EAA) ultrasound study on healthy, fertile men: scrotal ultrasound reference ranges and associations with clinical, seminal and biochemical characteristics. *Andrology*. 2020;9(2):559-576.
6. Lotti F, Frizza F, Balercia G, et al. The European Academy of Andrology (EAA) ultrasound study on healthy, fertile men: scrotal ultrasound reference ranges and associations with clinical, seminal, and biochemical characteristics. *Andrology*. 2021;9(2):559-576.
7. Forti G, Krausz C. Clinical review 100: evaluation and treatment of the infertile couple. *J Clin Endocrinol Metab*. 1998;83(12):4177-4188.
8. Spaggiari G, MG AR, Santi D. Testicular ultrasound inhomogeneity is an informative parameter for fertility evaluation. *Asian J Androl*. 2020;22(3):302-308.
9. Condorelli R, Calogero AE, La Vignera S. Relationship between testicular volume and conventional or nonconventional sperm parameters. *Int J Endocrinol*. 2013;2013:145792.
10. Fedder J. Prevalence of small testicular hyperechogenic foci in subgroups of 382 non-vasectomized, azoospermic men: a retrospective cohort study. *Andrology*. 2017;5(2):248-255.
11. Ventimiglia E, Ippolito S, Capogrosso P, et al. Primary, secondary and compensated hypogonadism: a novel risk stratification for infertile men. *Andrology*. 2017;5(3):505-510.
12. Cocuzza MS, Tiseo BC, Srougi V, et al. Diagnostic accuracy of physical examination compared with color Doppler ultrasound in the determination of varicocele diagnosis and grading: impact of urologists' experience. *Andrology*. 2020;8(5):1160-1166.
13. D'Andrea S, Martorella A, Castellini C, et al. Clinical and seminal parameters associated with testicular microlithiasis and its severity in males from infertile couples. *Hum Reprod*. 2021;36(4):891-898.
14. Isidori AM, Cantisani V, Giannetta E, et al. Multiparametric ultrasonography and ultrasound elastography in the differentiation of parathyroid lesions from ectopic thyroid lesions or lymphadenopathies. *Endocrine*. 2017;57(2):335-343.
15. Loberant N, Bhatt S, McLennan GT, Dogra VS. Striated appearance of the testes. *Ultrasound Q*. 2010;26(1):37-44.
16. Pozza C, Kanakis G, Carlomagno F, et al. Testicular ultrasound score: a new proposal for a scoring system to predict testicular function. *Andrology*. 2020;8(5):1051-1063.
17. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer Oxford Engl 1990)*. 2012;48(4):441-446.
18. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563-577.
19. WHO *WHO Laboratory Manual for the Examination and Processing of Human Semen*. 5th ed. WHO. 2010.
20. Sakamoto H, Ogawa Y. Does a clinical varicocele influence the relationship between testicular volume by ultrasound and testicular function in patients with infertility?. *Fertil Sterility*. 2009;92(5):1632-1637.
21. Sakamoto H, Yajima T, Nagata M, Okumura T, Suzuki K, Ogawa Y. Relationship between testicular size by ultrasonography and testicular function: measurement of testicular length, width, and depth in patients with infertility. *Int J Urol Off J Jpn Urol Assoc*. 2008;15(6):529-533.
22. Lenz S, Giwercman A, Elsborg A, et al. Ultrasonic testicular texture and size in 444 men from the general population: correlation to semen quality. *Eur Urol*. 1993;24(2):231-238.
23. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern*. 1973;SMC-3(6):610-621.

24. Galloway MM. Texture analysis using gray level run lengths. *Comput Graphics Image Process*. 1975;4(2):172-179.
25. Thibault G, Fertil B, Navarro C, et al. Texture indexes and gray level size zone matrix application to cell nuclei classification. *10th International Conference on Pattern Recognition and Information Processing, PRIP 2009*, 2009, Minsk, Blarus. 140–145.
26. Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern*. 1989;19(5):1264-1274.
27. Ruiz-Olvera SF, Rajmil O, Sanchez-Curbelo JR, Vinay J, Rodriguez-Espinosa J, Ruiz-Castane E. Association of serum testosterone levels and testicular volume in adult patients. *Andrologia*. 2018;50(3). https://doi/org/10.1111/and.12933
28. Einstein DM, Paushter DM, Singer AA, Thomas AJ, Levin HS. Fibrotic lesions of the testicle: sonographic patterns mimicking malignancy. *Urol Radiol*. 1992;14(3):205-210.
29. Harris RD, Chouteau C, Partrick M, Schned A. Prevalence and significance of heterogeneous testes revealed on sonography: ex vivo sonographic-pathologic correlation. *AJR Am J Roentgenol*. 2000;175(2):347-352.
30. Lenz S, Thomsen JK, Giwercman A, Hertel NT, Hertz J, Skakkebaek NE. Ultrasonic texture and volume of testicles in infertile men. *Hum Reprod Oxford Engl*. 1994;9(5):878-881.
31. Lenz S, Skakkebaek NE, Hertel NT. Abnormal ultrasonic pattern in contralateral testes in patients with unilateral testicular cancer. *World J Urol*. 1996;14(Suppl 1):S55-S58.
32. Yoon J, Lee E, Kang SW, Han K, Park VY, Kwak JY. Implications of US radiomics signature for predicting malignancy in thyroid nodules with indeterminate cytology. *Eur Radiol*. 2021;31(7):5059-5067.
33. Park KW, Shin JH, Hahn SY, Kim JH, Lim Y, Choi JY. The role of histogram analysis of grayscale sonograms to differentiate thyroid nodules identified by 18F-FDG PET-CT. *Medicine*. 2020;99(48):e23252.
34. Paris MT, Mourtzakis M. Muscle composition analysis of ultrasound images: a narrative review of texture analysis. *Ultrasound Med Biol*. 2021;47:880-895.
35. Satoh Y, Hirata K, Tamada D, Funayama S, Onishi H. Texture analysis in the diagnosis of primary breast cancer: comparison of high-resolution dedicated breast positron emission tomography (dbPET) and whole-body PET/CT. *Front Med*. 2020;7:603303.
36. Osapoetra LO, Chan W, Tran W, Kolios MC, Czarnota GJ. Comparison of methods for texture analysis of QUS parametric images in the characterization of breast lesions. *PLoS One*. 2020;15(12):e0244965.
37. Ardakani AA, Sattar A, Abolghasemi J, Mohammadi A. Correlation between kidney function and sonographic texture features after allograft transplantation with corresponding to serum creatinine: a long term follow-up study. *J Biomed Phys Eng*. 2020;10(6):713-726.
38. Huang X, Chen M, Liu P, Du Y. Texture feature-based classification on transrectal ultrasound image for prostatic cancer detection. *Comput Math Methods Med*. 2020;2020:7359375.
39. Miller WL, Auchus RJ. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocr Rev*. 2011;32(1):81-151.
40. Medar MLJ, Marinkovic DZ, Kojic Z, et al. Dependence of Leydig cell's mitochondrial physiology on luteinizing hormone signaling. *Life Basel*. 2020;11(1):19.
41. Rohayem J, Zitzmann M, Laurentino S, et al. The role of gonadotropins in testicular and adrenal androgen biosynthesis pathways—insights from males with congenital hypogonadotropic hypogonadism on hCG/rFSH and on testosterone replacement. *Clin Endocrinol*. 2021;94(1):90-101.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: De Santi B, Spaggiari G, Granata ARM, et al. From subjective to objective: a pilot study on testicular radiomics analysis as a measure of gonadal function. *Andrology*. 2022;10:505–517. https://doi.org/10.1111/andr.13131