Fournier CS, et al., J Reprod Med Gynecol Obstet 2024, 9: 174
DOI: 10.24966/RMGO-2574/100174

HSOA Journal of
Reproductive Medicine, Gynecology & Obstetrics

Research Article

Retrospective Study Documents Improvement in Semen Quality in Recent Years and AI Models Predict the Critical Parameters

Fournier CS[1,2]*, Iqbal SMA[3], Asghar W[1,4] and Matilsky M[2,4]

[1]*Department of Biological Sciences, Florida Atlantic University, Boca Raton, USA*

[2]*Boca Fertility, Boca Raton, USA*

[3]*Department of Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, USA*

[4]*Department of Biological Sciences, Florida Atlantic University, Boca Raton, FL, USA*

Abstract

The purpose of this retrospective study was to investigate the trend of semen quality over time and develop Artificial Intelligence (AI) models to identify parameters that may determine the quality of semen samples. We combined 10,480 semen analyses carried out at one fertility clinic in Florida, USA over 24 years, using the same methodology and under the supervision of one laboratory director. We conducted one-way ANOVA and post-hoc comparison to compare Concentration, Total Count, Motility, Total Motile Count, Forward Progression, and Morphology of spermatozoa from 2000 to 2023. We then created regression and classification AI models to determine which sperm parameters could predict and classify semen quality. We found a significant decline over time for Motility and Morphology. However, despite a significant decline from 2000-2003 to 2016-2019 for Concentration, Total Count, Total Motile Count, and Forward Progression, we found a significant increase from 2016-2019 to 2020-2023. We also found that the Volume and Concentration of samples were most effective in predicting Motility, Morphology and Forward Progression of spermatozoa.

Decline in sperm quality has been observed worldwide and male infertility is now known to be important in the need for Assisted Reproductive Technology procedures. However, the life-long production of spermatozoa makes the improvement of sperm quality a possibility with lifestyle adjustments. Although AI techniques have revolutionized dataset analysis in many scientific fields, little use of AI has been observed in andrology. Investigating sperm quality has major implications for helping patients going through fertility treatment and finding ways to increase their chances of success.

Keywords: Artificial intelligence; Improved semen quality; Male infertility; Semen analysis

Highlights

- Increase of sperm Concentration, Total Count, Total Motile Count, and Forward Progression from 2016 to 2023
- Developed AI predictive models to determine critical semen parameters

Introduction

Human infertility, defined as a failure to establish a clinical pregnancy after 12 months of regular, unprotected sexual intercourse [1], has been rapidly increasing worldwide in the past decades and affects more than 15% of couples [2]. Previously thought to be attributed to societal factors, such as industrialization or women's participation in the workforce, it is now clear that infertility is associated with biological parameters as well [3]. Assessing the causes of patient infertility is necessary to help increase success rates and achieve pregnancies through Assisted Reproductive Technologies (ART). ART is defined as any clinical intervention that includes the in-vitro handling of human oocytes, spermatozoa, or embryos for the purpose of reproduction [1]. One of the major factors predicting successful fertilization and ART outcomes is the use of normal, healthy spermatozoa.

Male infertility and reproductive health have largely been overlooked for decades. The focus of infertility research for Reproductive Endocrinology and Infertility (REI) physicians has been on treating female patients [4]. However, the high demand for fertility treatment can be explained by an increased number of couples with male infertility factors, ranging from 20-70% [5,6]. Male factors encompass abnormal semen parameters or function, anatomical, endocrine, genetic, functional or immunological abnormalities of the reproductive system, chronic illness, or sexual conditions incompatible with the ability to deposit semen in the vagina [1,3]. Understanding sperm quality has major implications in helping patients going through natural conception or fertility treatment because spermatozoa exhibiting optimal parameters improve embryo quality, pregnancy rate and live birth rate. An evaluation of semen and sperm parameters is needed prior to engaging in any clinical treatment of infertility. The World Health Organization (WHO) has published six manuals since 1980 to standardize procedures and methods for the evaluation of human semen. WHO also publishes the lower reference limit for normality around the world [7-9]. The 2nd to 4th editions' lower limits were empirical reference values, whereas the 5th and 6th editions' were based on published population-based analyses using 1,953 patients from eight countries on four continents and using 3,589 patients from 12 countries on five continents respectively [10,11]. The lower reference

---

***Corresponding author:** Fournier CS, Department of Biological Sciences, Florida Atlantic University, 777 Glades Road, Boca Raton, 33431, USA, Tel: +1 3058777478; Email: cfournier2015@fau.edu

Citation: Fournier CS, Iqbal SMA, Asghar W, Matilsky M (2024) Retrospective Study Documents Improvement in Semen Quality in Recent Years and AI Models Predict the Critical Parameters. J Reprod Med Gynecol Obstet 9: 174.

Received: September 07, 2024; **Accepted:** September 18, 2024; **Published:** September 25, 2024

Year	# analyses	Age (years)	Abstinence (days)	pH	Volume (ml)	Concentration(million/ml)	Total Count (million)	Motility (%)	Total Motile Count (million)	Forward Progression (%)	Morphology (% Normal)
2000	142	37.07	4.10	8.2	3.12	73.85	203.15	58.42	143.21	46.75	25.72
2001	183	38.42	3.82	8.2	2.86	69.65	199.41	59.72	145.25	47.82	27.68
2002	180	36.81	3.91	8.2	3.16	70.80	207.96	65.06	157.98	54.31	29.44
2003	216	37.32	4.00	8.2	2.91	63.81	169.58	62.60	129.74	49.96	28.60
2004	201	37.17	3.88	8.3	2.89	83.96	223.28	67.03	176.53	54.77	17.88
2005	244	38.25	4.00	8.2	2.98	64.22	188.30	64.19	144.47	51.81	10.20
2006	241	37.76	4.02	8.1	3.01	57.73	172.33	60.27	132.60	51.04	10.39
2007	206	37.18	3.80	8.1	2.68	71.88	175.44	60.95	127.02	50.74	10.66
2008	227	37.78	3.95	8.1	3.09	74.91	210.66	61.42	157.99	52.19	11.50
2009	213	37.80	4.01	8.1	2.87	62.63	174.76	63.11	127.49	50.37	8.96
2010	270	37.81	4.14	8.1	2.97	68.73	196.83	60.45	140.17	48.79	8.93
2011	694	36.64	4.21	8.0	2.53	50.32	126.25	54.99	89.17	45.22	5.90
2012	463	36.84	3.97	8.0	2.69	58.56	154.29	57.32	113.09	44.47	5.72
2013	378	36.54	3.91	8.0	2.63	58.74	149.70	58.31	107.37	47.48	5.18
2014	399	38.77	3.93	7.9	2.64	54.66	145.58	55.13	102.51	43.98	4.44
2015	402	36.40	3.90	7.7	2.86	49.47	136.48	52.29	93.76	38.85	3.95
2016	486	36.67	3.92	7.6	2.66	49.56	127.96	53.25	86.71	43.75	3.86
2017	517	37.32	3.96	7.7	2.82	48.57	130.18	53.81	87.13	40.13	4.37
2018	573	37.43	4.01	7.7	2.69	47.56	129.30	51.86	89.88	37.84	5.02
2019	534	38.08	4.22	7.6	2.81	57.88	157.93	51.61	102.64	38.61	4.65
2020	656	36.86	4.15	7.7	2.87	59.64	169.13	51.47	114.30	38.48	4.83
2021	1,022	37.33	4.31	7.8	2.92	65.06	181.16	52.20	123.68	43.10	3.68
2022	1,048	38.20	4.28	7.8	2.80	64.51	168.31	49.53	111.85	42.96	3.36
2023	985	37.71	4.12	7.9	2.78	63.54	161.81	48.97	110.27	37.94	2.66
Total	**10,480**	**37.43**	**4.08**	**7.9**	**2.80**	**60.02**	**161.69**	**54.60**	**113.24**	**43.59**	**7.00**

Table 1: Average per year of different semen parameters collected at the Boca Fertility clinic between 2000 and 2023.

Average semen parameters per year

limits are calculated based on the 5th centile of the combined published data [10,11]. Data show a decrease in the lower reference limit for volume, concentration, motility, and morphology between 1999 and 2010, but a slight increase in 2021 for concentration and motility [7-12].

Artificial Intelligence (AI) and Machine Learning (ML) techniques have revolutionized dataset analysis in many scientific fields; however, little use of AI & ML has been observed in andrology [13]. Some research has used these techniques on processed semen samples by analyzing samples under optical microscopy, classifying normal and abnormal sperm samples, and identifying factors responsible for affecting sperm quality. For instance, Agarwal et al. used AI Optical Microscopic (AIOM)-based technology to measure sperm motility, concentration, and pH. They found a high correlation between the AI-based technique and the standard manual method (N = 135 samples). For concentration and motility, the Spearman's rank correlation coefficients (r) were found to be highly correlated (0.974 and 0.925 respectively) [14]. Similarly, Butola et al., used Feed-Forward Deep Neural Networks (FFDNN) to differentiate normal semen samples from samples that were under stress conditions, such as cryopreservation or exposure to ethanol or hydrogen peroxide [15]. They developed a Spatially Coherent Digital Holographic Microscope (PSC-DHM) for the quantitative phase imaging of sperm samples [15]. They used these phase maps in developing seven FFDNN models for classifying normal sperm samples from those under stress, and found FFDNN model to be 85.6% accurate, 85.5% sensitive, and 94.7% specific. In addition, Wu et al., used logistic regression to identify factors that affect sperm quality by studying 5042 ejaculates from 385 boars [16]. They found that Fe deficiency, excess Cu, and the presence of Pb were factors involved in the reduced quality of sperm [16]. Similarly, Ory et al., used Random Forest (RF), Logistic Regression (LR), and Support Vector Machine (SVM) to predict the upgradation of sperm parameters after Varicocele Repair [17]. They found that 45.6% of men (N = 45) experienced improvement in their sperm concentration and the RF model was found to be accurate in making the prediction [17].

A major drawback from review publications on sperm quality is the combining of studies from different countries and clinics, using different methodologies, technicians, personnel experience, and timeframes. Here, we focused the investigation on sperm quality through a retrospective study from semen analyses collected at a single clinic, with consistent methods and trained personnel under the supervision of one laboratory director. The purpose of this study was to use semen analysis data routinely collected at a fertility clinic to investigate sperm quality over time. We hypothesized that we would observe sperm quality decline for each parameter investigated and a decline in normozoospermic samples similar to data recorded worldwide. We also used our comprehensive dataset to develop AI models (regression and classification models) and identify features responsible for predicting significant factors that play a role in determining the quality of semen samples.

Results

Sperm quality over time

Data used for analysis are summarized in table 1, where samples per year have been combined to yield an average for all parameters. The mean, standard deviation, standard error, and range of data for the patients' age, number of days of abstinence, pH, semen volume, and the six sperm parameters analyzed are recorded in table 2.

	Mean	SD	SE	Range
Age (years)	37.43	7.60	0.075	13-76
Abstinence (days)	4.09	2.50	0.025	0-120
pH	7.87	0.37	0.004	4-10
Volume (ml)	2.80	1.56	0.015	0.1-14.8
Concentration (106/ml)	60.73	63.73	0.623	0-400
Total Count (106)	161.69	195.76	1.91	0-2,000
Motility (%)	54.60	27.75	0.271	0-100
Total Motile Count (x106)	113.25	153.66	1.501	0-1,425
Forward Progression (%)	43.67	26.91	0.270	0-95
Morphology (% Normal)	7.00	9.39	0.095	0-65

Table 2: Mean, Standard Deviation (SD), Standard Error (SE), and range of different parameters collected from 10,480 semen analyses from 2000-2023 at the Boca Fertility clinic.

Mean, SD, SE, and range of different semen parameters

We observed a statistically significant difference in sperm quality for the following parameters: Concentration (F = 18.99, df = 5, $p < 0.001$), Total Count (F = 17.71, df = 5, $p < 0.001$), Motility (F = 17.71, df = 5, $p < 0.001$), Total Motile Count (F = 22.72, df = 5, $p < 0.001$), Forward Progression (F = 46.38, df = 5, $p < 0.001$), and Morphology (F = 38.79, df = 5, $p < 0.001$) between four years' increments (Figure 1).

Post-hoc comparisons with the Tukey HSD tests showed a significant decrease ($p < 0.001$) from 2000-2003 and 2016-2019 for all six parameters evaluated (Figure 1). Post-hoc comparisons also showed a significant decrease from years 2016-2019 and 2020-2023 for Motility ($p < 0.05$) and Morphology ($p < 0.001$; Figure 1). However, they showed a significant increase ($p < 0.001$) for Concentration, Total Count, Total Motile Count and a non-significant increase for Forward Progression ($p = 0.857$; Figure 1).

It is important to note that despite a significant decline throughout the years, means of Concentration, Motility, and Forward Progression per year were still within normal range (Table 1) [9]. However, the proportion of normozoospermic samples declined throughout the years and the proportion of samples with four abnormal parameters (Concentration, Motility, Forward Progression and Morphology) increased (Figure 2).

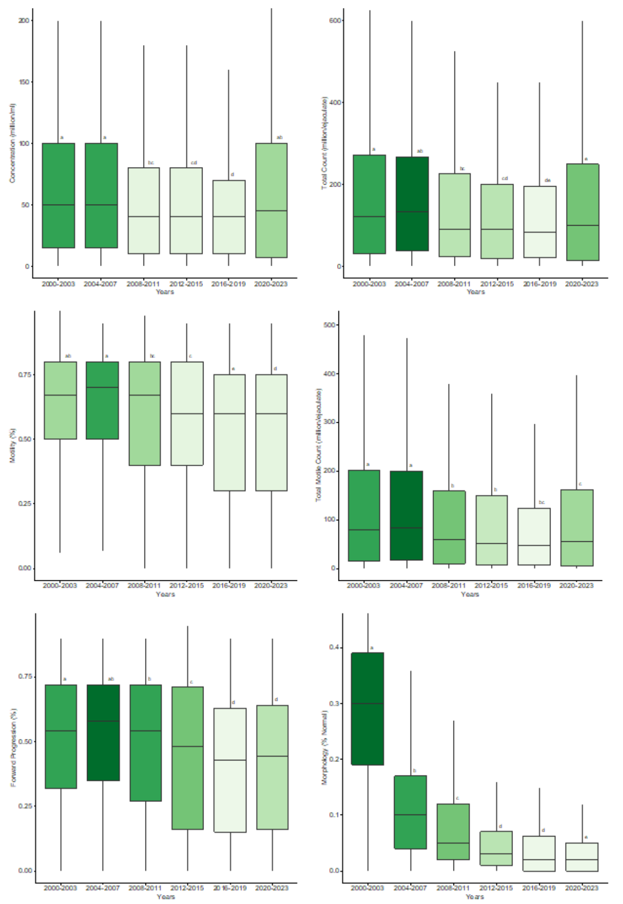

Figure 1: Box plot of the six sperm parameters evaluated between 2000-2023. Different letters represent significant difference at a $p < 0.05$ probability level.

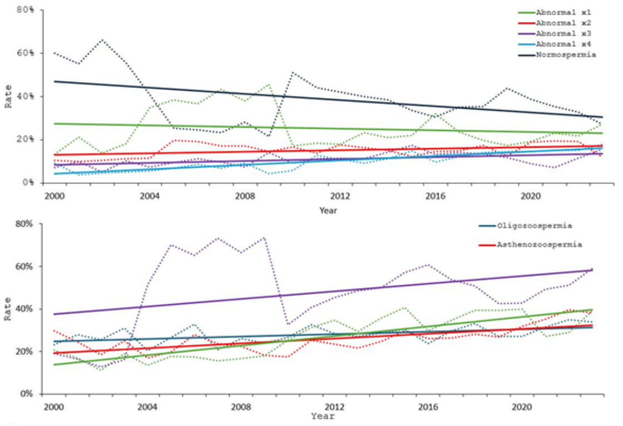

Figure 2: Top. Rate of semen samples with abnormal parameters based on the total number of samples per year between 2000 and 2023. Bottom. Rate of oligo-, astheno-, teratozoospermic, abnormal forward progression, and normospermic samples based on the total number of samples per year between 2000 and 2023.

Moreover, observed and expected counts of Normal and Abnormal samples (containing one, two, three, or four abnormal parameters) for Concentration, Motility, Forward Progression, and Morphology were statistically significant ($\chi^2 = 371.19$, df = 20, $p < 0.001$; Figure 3). Normozoospermic samples were observed more frequently than expected by chance in 2000-2003, but less frequently than expected by chance in 2020-2023 (Figure 3).

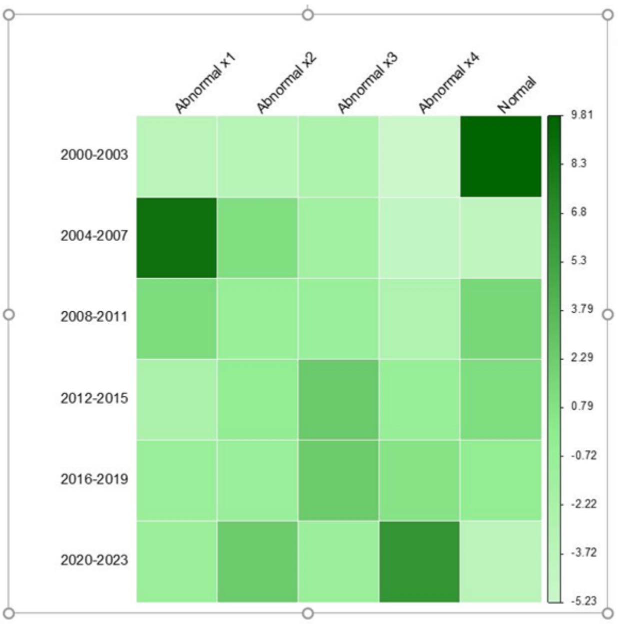

Figure 3: Plot of matrix for a Pearson's Chi-square test residual values ($\chi^2 = 371.19$, df = 20, $p < 0.001$) for Normal and Abnormal samples (containing one, two, three, or four abnormal parameters) for Concentration, Motility, Forward Progression, and Morphology throughout the years.

Artificial Intelligence Models

Regression models

Multivariate techniques have been used for modeling regression models, which allows prediction of multiple responses from the use of multiple features. Different supervised learning regression models, such as Random Forest (RF), Linear Regression (LR), and Support Vector Regression (SVR) models were applied to predict parameters using different combination of features and responses. The widely used non-linear RBF kernel was used for SVR. We also used data transformation techniques to optimize these models. The aim was to determine parameters that were significant in predicting the quality of a semen sample. Based on this, the initial model was:

Responses ~ Features

Morphology + Motility + Forward Progression + pH ~ Age + Year + Volume + Concentration (Model 1)

We applied RF, LR, and SVR to evaluate the predictive ability of Model 1. From the results, we optimized it by using data transformation techniques and by removing parameters that were less significant and responsible for reducing the overall predictive ability of the model. Models were evaluated based on their R-squared values (R^2) using standard 10-fold cross validation.

The mean R^2 values of Model 1 using RF, LR, and SVR were 65.2%, 49.9%, and 2.2% respectively. From these values, RF outperformed the other two models. However, 65.2% was not a significant predictive ability, as the best regressor still failed to predict almost 35% of the data. Model 1 was further analyzed by splitting the entire dataset into 80-20% (training and test datasets) and plotting the predicted vs actual value plots for all the responses (Figure 4). Of the four responses, pH did not have a predictive trend and hence did not predict Model 1 well.

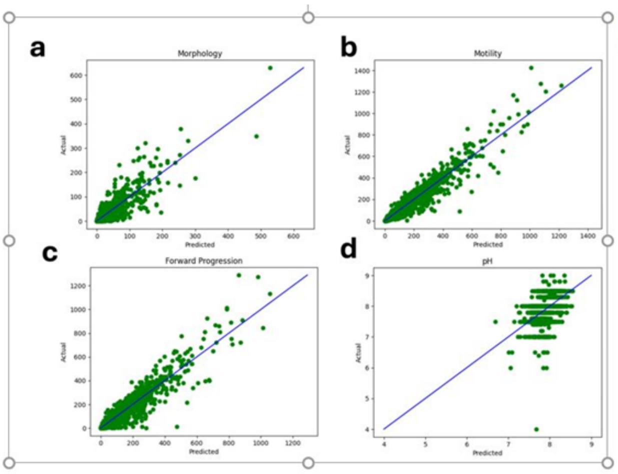

Figure 4: Actual vs predicted values for Model 1 for a) Normal Morphology, b) Motility, c) Forward Progression, and d) pH.

Therefore, we dropped pH from Model 1 and ran RF, LR, and SVR on Model 2:

Morphology + Motility + Forward Progression ~ Age + Year + Volume + Concentration (Model 2)

The mean R^2 for RF, LR, and SVR were 83.5%, 61.2%, and 2.0% respectively. The mean R^2 for RF increased by 28% from Model 1 to Model 2. To study the effect of Volume and Concentration on Morphology, Motility, and Forward Progression, we developed Model 3 with Volume and Concentration as features and Morphology, Motility, and Forward Progression as responses, and applied the previously mentioned regressors:

Morphology + Motility + Forward Progression ~ Volume + Concentration (Model 3)

The mean R^2 for RF, LR, and SVR were 74.9%, 57.7%, and 40.8% respectively. RF once again performed considerably better than the other two regressors; however, it is highly desirable to increase the predictive ability of the regressors. For this purpose, different data transformation techniques are generally used to increase the performance of the models, including square root, logarithmic, and cubic transformations. Data transformation helps in organizing the data that eventually leads to better accuracy of the model [18]. We tried different transformations for our dataset and found encouraging results for the square root transformation. Therefore, we transformed Model 3 using square root data transformation to form Model 4:

$\text{Morphology}^{0.5} + \text{Motility}^{0.5} + \text{Forward Progression}^{0.5} \sim$ Volume + Concentration (Model 4)

The mean R^2 for RF, LR, and SVR were 82.0%, 68.9%, and 71.2% respectively. With data transformation, the R^2 value using RF regressor increased by 9.5%. It is worth noting that Model 4 utilizes only two parameters as features and predicts three responses with high accuracy and hence, highlights the efficacy of the model for predicting useful information about the sample with limited information available. Figure 5 shows the actual vs predicted values for Morphology, Motility, and Forward Progression using Volume and Concentration of samples. The minimum R^2 for any fold in Model 4 using RF was 80.8%. An 82.0% predictive ability means the model can predict

82.0% of the instances and every fold with R^2 greater than 80.8% means that the model using RF was not over fitted to any part of the dataset but rather can be generalized to the entire dataset.

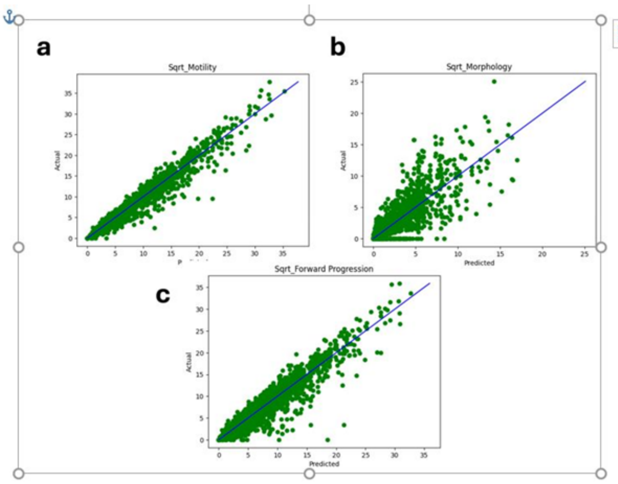

Figure 5: Actual vs predicted values for Model 4 for a) Normal Morphology$^{0.5}$, b) Motility$^{0.5}$, and c) Forward Progression$^{0.5}$.

Classification models

Along with regression models, we were also interested in classifying significant responses into Normal and Abnormal, which could be useful in andrology to highlight different sperm parameters. For this purpose, we applied classification techniques, such as RF and Decision Tree (DT)classification techniques, such as RF and Decision Tree (DT) using our dataset to classify parameters that can be significant towards predicting the quality of semen samples. We categorized the parameters of interest, Morphology, Motility, and Forward Progression as Normal or Abnormal based on the thresholds set by WHO for these parameters throughout the years. We used both temporal and non-temporal parameters to classify Motility, Morphology, and Forward Progression. Model 5 represents the first classification model used:

Motility ~ Age + Year + Volume + Concentration + pH (Model 5a)

Morphology ~ Age + Year + Volume + Concentration + pH (Model 5b)

Forward Progression ~ Age + Year + Volume + Concentration + pH (Model 5b)

We used optimized hyperparameters for modeling the above models. These hyperparameters include maximum depth of the tree, minimum number of sample split, and minimum sample leaf. We then applied 10-fold cross validation using the tuner parameters. Decision Tree (DT) classifies Motility into Normal and Abnormal with the highest accuracy of 72.5%, classifies Morphology with an accuracy of 65.9% and Forward Progression with an accuracy of 75.9% (Table 3). RF classifier was found to classify parameters with higher accuracies than DT with 77.5%, 66.3% and 75.7% accuracies for Motility, Morphology, and Forward Progression respectively. We also trained RF and DT classification models for temporal and non-temporal models separately to identify their effectiveness in determining the abnormality of Motility, Morphology, and Forward Progression:

Motility ~ Age + Year (Model 6a)

Morphology ~ Age + Year (Model 6b)

Forward Progression ~ Age + Year (Model 6c)

Motility ~ Volume + Concentration (Model 7a)

Morphology ~ Volume + Concentration (Model 7b)

Forward Progression ~ Volume + Concentration (Model 7c)

	Decision Tree			**Random Forest**		
Parameters	Max Depth of Tree	Min Sample Leaf	Min Sample Split	Max Depth of Tree	Min Sample Leaf	Min Sample Split
Model 5a	5	1	2	5	3	4
Model 5b	5	1	2	5	4	6
Model 5c	5	4	2	5	3	8
Model 6a	50	3	4	5	2	4
Model 6b	10	2	2	5	1	6
Model 6c	5	2	2	5	2	4
Model 7a	5	3	2	5	1	6
Model 7b	10	3	2	5	3	6
Model 7c	5	2	2	5	1	8

Table 3: Best parameters for classifying motility and morphology using decision tree and random forest.

Summary of parameters using Decision Tree and Random Forest

In Models 6a-c, DT classified Motility, Morphology, and Forward Progression with an accuracy of 30.9%, 39.9%, and 32.2% respectively, whereas RF classified them with 59.5%, 46.3%, and 35.9% respectively. Similarly, in Models 7a-c, DT classified Motility with accuracy of 83.9%, Morphology with 69.9% and Forward Progression with 79.6%, whereas RF classified Motility with 84.2%, Morphology with 70.8%, and Forward Progression with 79.9% accuracy. Both DT and RF models classified Motility, Morphology, and Forward Progression using Volume and Concentration with higher accuracies than Age and Year. These results show that temporal parameters were not suitable for classifying Motility and Morphology as Normal and Abnormal and overall, RF is a better classifier than a simple Decision Tree.

Discussion

Results from our 24-yearlong comprehensive dataset highlighted an improvement in sperm quality from 2020-2023 for four of the six parameters examined, even though a significant decline was observed since 2000 for Motility and Morphology. Normal morphology of spermatozoa is known to be associated with high progressive motility. Studies examining sperm morphology at three different levels of the female reproductive tract [19-28] show that more than 90% of the morphologically normal cells are found at the highest level, in the endocervical mucus [29]. A small increase in Concentration and Motility parameters was observed since 2010 in the WHO 6th edition, from data collected worldwide [10,11]. Similar results regarding Concentration, Total Count, Total Motile Count or velocity parameters emerged in the literature recently, involving worldwide studies. Infertility issues have come to be widely acknowledged by the general population. Couples trying to conceive are aware of their fertility status and try to improve their chances of success. In our dataset, the Covid 19 epidemic may have played a role in increasing patients' awareness of their health. As virtual work became more common, people may have increased free time to acknowledge the benefits of exercising and following a healthy diet while trying to expand their families. Continuous spermatogenesis makes possible the improvement of sperm quality with lifestyle adjustments. Most men are able to improve quality of most of their sperm parameters over time with supplements, better diet, or exercise [22-27].

Part of the results from this study were similar to those of worldwide studies [10,11,30-32]. Western, industrialized countries have been witnessing a major decline in sperm motility over the past 100 years [30,31,33,34]. Reasons behind this decline is uncertain but evidence suggest that modern lifestyle and environmental exposure, such as endocrine disruptors, poor diet, or low exercise, play a major role in male infertility and sperm quality decline [3,32,35-37]. Couples with male infertility factor, up to 70% in some areas, is one of the leading causes of increased demand for fertility treatment [5,6].

Along with statistical analysis, we used our dataset to develop AI models and determine parameters that are significant in predicting the overall quality of semen samples. To the best of our knowledge, no AI predictive models have yet been used in andrology. We found that of non-linear regression models, RF was the most predictive in modeling our dataset. We optimized these models using data transformation techniques and found square root transformation to be the most adequate in modeling the relationship between features and responses. In regression models, we found that semen Volume and sperm Concentration were most effective in predicting Motility, Morphology, and Forward Progression of spermatozoa. Similarly, we also used classification models to classify Motility, Morphology, and Forward Progression into Normal and Abnormal categories using DT and RF. In these models, we found similar results, which shows that these models were effective in determining relationships between parameters related to the quality of semen samples and were a good application of AI-based analysis in the andrology field.

One possible limitation of this study is that the population tested may not be representative of the whole American population as patients tested were conscious of their fertility status and many were about to proceed with an IVF cycle for infertility issues. Even though results come from only one relatively small fertility clinic, it is important to note that South Florida represents a melting-pot of patients from different origins, background, age, sexual orientation, or family composition. Even though the use of the Makler chamber to perform semen analyses does not strictly follow the guidelines described in the WHO manuals, this methodology has been widely used in fertility clinics. The consistent use of the Makler chamber over the 24 years of our study did not interfere with our interpretation of the results. In addition, our sample size has been steadily increasing since 2000, with an important increase since 2021, which demonstrates compelling results for the last 3-year increment.

In a future study, we would like to focus on patients that went through an IVF cycle at Boca Fertility and determine if semen parameters, at the day of retrieval, have an impact on the outcome of the cycle. We would investigate samples before and after processing and potentially correlate semen and sperm parameters to fertilization rate, embryo quality, and pregnancy rate. Using AI analyses, we could also attempt to predict different outcomes based on the semen sample quality used for the IVF cycle.

Conclusion

We used 24-year retrospective semen analysis data collected at one fertility clinic in Florida and observed a significant improvement in some of the sperm parameters investigated, despite a decline in normozoospermic samples over the time frame. We also used our dataset to develop regression and classification models and found encouraging results, which shows the applicability of different types of AI-based techniques in the field of andrology. Regression models gave us insight into different parameters of the sample and can help scientists better understand the quality of samples under consideration.

Materials and Methods

Semen analysis

Semen analyses were performed daily at Boca Fertility, a fertility clinic in Boca Raton, FL, USA, by embryologists/andrologists on either external patients or internal patients. External patients were referred by their urologists for investigation of potential anatomical or physiological problems or to preserve their fertility through sperm freezing for future use. Internal patients refer to clinical patients who were considering going through an *In Vitro* Fertilization (IVF) cycle. Semen analysis is the first step in investigating infertility and directing the future course of action for these patients. We recorded men's age, semen volume and pH, number of days of abstinence, sperm Concentration (million/ml), Motility (%), Forward Progression (%), and Morphology (% normal). We recorded Concentration, Motility, and Forward Progression using a Makler chamber using manufacturer's protocol (Sefi Medical, Israel) and Morphology using a smear stained with Papanicolaou stain [8] from 2000-2010 and pre-stained slides (Testsimplets®, Waldeck GmbH & Co. KG, Germany) from 2010-2023 [38]. Our clinic had been using the same semen analysis methodology since 2000, under the supervision of one single laboratory director, Dr. Michael Matilsky, PhD, HCLD, who trained all technologists involved in the routine semen analyses throughout the years.

Semen samples with low concentration of spermatozoa (oligozoospermia), reduced percentage of motile spermatozoa (asthenozoospermia), reduced percentage of progressively motile spermatozoa, and reduced percentage of morphologically normal spermatozoa (teratozoospermia) were determined when the numbers fell below the WHO lower reference limit for normality for the period studied [7-9].

Data analysis

In this study, we combined a total of 10,480 semen analyses carried out at the fertility clinic, over a 24-year period. For analysis purposes, we included parameters regularly recorded, such as Concentration (million/ml), Motility (%), Forward Progression (%), and Morphology (% normal), and we calculated Total Count (million/ejaculate), as Concentration multiplied by the ejaculate Volume, and Total Motile Count (million/ejaculate), as the Total Count multiplied by % Motility. We recorded the mean, standard deviation, standard error and range of values for all parameters.

All samples were collected by masturbation into a sterile cup, either at home or at the clinic. We excluded analyses from TESE (Testicular Sperm Extraction), MESA (Microsurgical Epididymal Sperm Aspiration), retrograde urine, post-vasectomy and post-vasectomy reversal samples or samples that were collected in a condom through sexual intercourse. We did not include analyses where the sample was collected for use in IVF or Intra-Uterine Insemination (IUI) cycles. On a regular basis, we did not record specific numbers for samples with extremely low concentration and labeled them as 'Rare' or 'Very Rare Motile' spermatozoa. For analysis purposes in this study, we determined Rare Motile spermatozoa as a concentration of 0.1 million/ml and a motility of 20% and Very Rare Motile spermatozoa as a concentration of 0.01 million/ml and a motility of 10%. No data on Forward Progression or Morphology were recorded for these samples.

From the WHO lower limits throughout the years [7-9], we categorized each sample as Normal or Abnormal based on their Concentration, Motility, Forward Progression, Morphology, and year of collection. We recorded the number of samples per year with one, two, three, or four abnormal parameters.

Statistical analysis

We combined data into increments of four years from 2000-2023 for analysis purposes. We conducted one-way ANOVA analyses and post-hoc comparisons using Tukey HSD tests to compare the six sperm parameters throughout the years. We also conducted a Pearson's Chi-squared ($\chi 2$) test of independence to compare counts of Normal and Abnormal samples (containing one, two, three, or four abnormal parameters) throughout the years. We ran all statistical analyses in R statistical software version 4.2.0.

Artificial intelligence models

We developed predictive and classification AI models to determine sperm parameters that were significant in predicting and classifying the quality of semen samples respectively. We performed a comparative analysis of these predictive models with different regression models and chose the models for better prediction of sperm quality. We also performed comparative analysis of classification models to identify the features that had better abilities to determine the parameters involved in higher sperm quality. We programmed these different models by using Python and Google Colab as an Integrated Development Environment (IDE).

Declarations

Ethics approval and consent to participate

All semen samples were voluntarily given by patients to Boca Fertility for clinical purposes. All patient-identifiable information were deleted before proceeding with this study. This research qualifies as exempt human subjects research under IRB regulations.

Availability of data and materials

The dataset used and analyzed during the current study is available from the corresponding author on reasonable request.

Conflict of interest

The authors declare that they have no competing interests.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author's contribution

CSF: Conceptualization, Data curation, Formal analysis of semen analysis data, Investigation, Methodology, Visualization, Writing - original draft;

SMAI: Formal analysis of AI models, Investigation, Methodology, Software, Visualization, Writing - original draft;

WA: Project administration, Supervision, Validation, Writing - review & editing;

MM: Conceptualization, Methodology, Project administration, Supervision, Validation, Writing - review & editing.

Acknowledgement

Not applicable.

References

1. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, et al. (2017) The International Glossary on Infertility and Fertility Care, 2017. Fertil Steril 108: 393-406.
2. Boivin J, Bunting L, Collins JA, Nygren KG (2007) International estimates of infertility prevalence and treatment-seeking: Potential need and demand for infertility medical care. Hum Reprod 22: 1506-1512.
3. Skakkebaek NE, Jørgensen N, Main KM, Rajpert-De Meyts E, Leffers H, et al. (2006) Is human fecundity declining? Int J Androl 29: 2-11.
4. De Jonge C, Barratt CLR (2019) The present crisis in male reproductive health: An urgent need for a political, social, and research roadmap. Andrology 7: 762-768.
5. Nosrati R, Graham PJ, Zhang B, Riordon J, Lagunov A, et al. (2017) Microfluidics for sperm analysis and selection. Nat Rev Urol 14: 707-730.
6. Agarwal A, Mulgund A, Hamada A, Chyatte MR (2015) A unique view on male infertility around the globe. Reprod Biol Endocrinol 13: 37.
7. World Health Organization (2010) WHO laboratory manual for the examination and processing of human semen, 5th ed. World Health Organization, Geneva, Switzerland.
8. World Health Organization (1999) WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction Fourth Edition. World Health Organization, Geneva, Switzerland.
9. World Health Organization (2021) WHO laboratory manual for the examination and processing of human semen (6thedn). World Health Organization, Geneva, Switzerland.
10. Campbell MJ, Lotti F, Baldi E, Schlatt S, Festin MPR, et al. (2021) Distribution of semen examination results 2020 - A follow up of data collated for the WHO semen analysis manual 2010. Andrology 9: 817-822.
11. Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HW, et al. (2010) World Health Organization reference values for human semen characteristics. Hum Reprod Updat 16: 231-45.
12. Wang C, Mbizvo M, Festin MP, Björndahl L, Toskin I (2022) Evolution of the WHO "Semen" processing manual from the first (1980) to the sixth edition (2021). Fertil Steril 117: 237-245.
13. Alsaidi M, Altaher A, Jan MT, Altaher ASA, Salekshahrezaee Z (2022) COVID-19 classification using deep learning two-stage approach.
14. Agarwal A, Henkel R, Huang CC, Lee MS (2019) Automation of human semen analysis using a novel artificial intelligence optical microscopic technology. Andrologia 51: 13440.
15. Butola A, Popova D, Prasad DK, Ahmad A, Habib A, et al. (2020) High spatially sensitive quantitative phase imaging assisted with deep neural network for classification of human spermatozoa under stressed condition. Sci Rep 10: 13118.
16. Wu Y, Wang C, Tan J, Wei HK, Sun H, et al. (2019) Logistic Regression Analysis Factors Affecting Sperm Motility and Abnormal Sperm Morphology in Boars. Animals (Basel) 9: 1004.

17. Ory J, Tradewell MB, Blankstein U, Lima TF, Nackeeran S, et al. (2022) Artificial Intelligence Based Machine Learning Models Predict Sperm Parameter Upgrading after Varicocele Repair: A Multi-Institutional Analysis. World J Mens Health 40: 618-626.

18. TIBCO (2024) What is data transformation? TIBCO, Santa Clara, USA.

19. Li Y, Lu T, Wu Z, Wang Z, Yu T, et al. (2023) Trends in sperm quality by computer-assisted sperm analysis of 49,189 men during 2015-2021 in a fertility center from China. Front Endocrinol (Lausanne) 14: 1194455.

20. Luo X, Yin C, Shi Y, Du C, Pan X (2023) Global trends in semen quality of young men: A systematic review and regression analysis. J Assist Reprod Genet 40: 1807-1816.

21. Lassen E, Pacey A, Skytte A-B, Montgomerie R (2024) Recent decline in sperm motility among donor candidates at a sperm bank in Denmark. Human Reproduction 39: 1618-1627.

22. Fett R (2023) It Starts with the Egg: The Science of Egg Quality for Fertility, Miscarriage, and IVF (3rdedn). Franklin Fox Publishing LLC Wayne, USA.

23. Schmid TE, Eskenazi B, Marchetti F, Young S, Weldon RH, et al. (2012) Micronutrients intake is associated with improved sperm DNA quality in older men. Fertil Steril 98: 1130-1137.

24. Salas-Huetos A, James ER, Aston KI, Jenkins TG, Carrell DT (2019) Diet and sperm quality: Nutrients, foods and dietary patterns. Reprod Biol 19: 219-224.

25. Cohen N, Ben-Meir A, Harlap T, Imbar T, Karavani G (2024) Changes in Sperm Parameters with Time in Men with Normal and Abnormal Baseline Semen Analysis. Reprod Sci 31: 1712-1718.

26. Sharma A, Papanikolaou N, Abou Sherif S, Dimakopolou A, Thaventhiran T, et al. (2024) Improvements in Sperm Motility Following Low- or High-Intensity Dietary Interventions in Men With Obesity. J Clin Endocrinol Metab 109: 449-460.

27. Sun B, Messerlian C, Sun ZH, Duan P, Chen HG, et al. (2019) Physical activity and sedentary time in relation to semen quality in healthy men screened as potential sperm donors. Hum Reprod 34: 2330-2339.

28. van Zyl JA (1980) The infertile couple. Part II. Examination and evaluation of semen. S Afr Med J 57: 485-491.

29. Mortimer D, Menkveld R (2001) Sperm morphology assessment--historical perspectives and current opinions. J Androl 22: 192-204.

30. Levine H, Jørgensen N, Martino-Andrade A, Mendiola J, Weksler-Derri D, et al. (2017) Temporal trends in sperm count: A systematic review and meta-regression analysis. Hum Reprod Update 23: 646-659.

31. Levine H, Jørgensen N, Martino-Andrade A, Mendiola J, Weksler-Derri D, et al. (2023) Temporal trends in sperm count: A systematic review and meta-regression analysis of samples collected globally in the 20th and 21st centuries. Hum Reprod Update 29: 157-176.

32. Swan SH, Colino S (2020) Count down: How our modern world is threatening sperm counts, altering male and female reproductive development, and imperiling the future of the human race. Scribner, New York, USA.

33. Carlsen E, Giwercman A, Keiding N, Skakkebaek NE (1992) Evidence for decreasing quality of semen during past 50 years. BMJ 305: 609-613.

34. Centola GM, Blanchard A, Demick J, Li S, Eisenberg ML (2016) Decline in sperm count and motility in young adult men from 2003 to 2013: Observations from a U.S. sperm bank. Andrology 4: 270-276.

35. Wang H, He H, Wei Y, Gao X, Zhang T, et al. (2023) Do phthalates and their metabolites cause poor semen quality? A systematic review and meta-analysis of epidemiological studies on risk of decline in sperm quality. Environ Sci Pollut Res Int 30: 34214-34228.

36. Skakkebæk NE, Lindahl-Jacobsen R, Levine H, Andersson AM, Jørgensen N, et al. (2022) Environmental factors in declining human fertility. Nat Rev Endocrinol 18: 139-157.

37. Virtanen HE, Jørgensen N, Toppari J (2017) Semen quality in the 21st century. Nat Rev Urol 14: 120-130.

38. Schirren C, Eckhardt U, Jachczik R, Carstensen CA (1977) Morphological differentiation of human spermatozoa with Testsimplets slides. Andrologia 9: 191-192.

HERALD

Advances In Industrial Biotechnology | ISSN: 2639-5665

Advances In Microbiology Research | ISSN: 2689-694X

Archives Of Surgery And Surgical Education | ISSN: 2689-3126

Archives Of Urology

Archives Of Zoological Studies | ISSN: 2640-7779

Current Trends Medical And Biological Engineering

International Journal Of Case Reports And Therapeutic Studies | ISSN: 2689-310X

Journal Of Addiction & Addictive Disorders | ISSN: 2578-7276

Journal Of Agronomy & Agricultural Science | ISSN: 2689-8292

Journal Of AIDS Clinical Research & STDs | ISSN: 2572-7370

Journal Of Alcoholism Drug Abuse & Substance Dependence | ISSN: 2572-9594

Journal Of Allergy Disorders & Therapy | ISSN: 2470-749X

Journal Of Alternative Complementary & Integrative Medicine | ISSN: 2470-7562

Journal Of Alzheimers & Neurodegenerative Diseases | ISSN: 2572-9608

Journal Of Anesthesia & Clinical Care | ISSN: 2378-8879

Journal Of Angiology & Vascular Surgery | ISSN: 2572-7397

Journal Of Animal Research & Veterinary Science | ISSN: 2639-3751

Journal Of Aquaculture & Fisheries | ISSN: 2576-5523

Journal Of Atmospheric & Earth Sciences | ISSN: 2689-8780

Journal Of Biotech Research & Biochemistry

Journal Of Brain & Neuroscience Research

Journal Of Cancer Biology & Treatment | ISSN: 2470-7546

Journal Of Cardiology Study & Research | ISSN: 2640-768X

Journal Of Cell Biology & Cell Metabolism | ISSN: 2381-1943

Journal Of Clinical Dermatology & Therapy | ISSN: 2378-8771

Journal Of Clinical Immunology & Immunotherapy | ISSN: 2378-8844

Journal Of Clinical Studies & Medical Case Reports | ISSN: 2378-8801

Journal Of Community Medicine & Public Health Care | ISSN: 2381-1978

Journal Of Cytology & Tissue Biology | ISSN: 2378-9107

Journal Of Dairy Research & Technology | ISSN: 2688-9315

Journal Of Dentistry Oral Health & Cosmesis | ISSN: 2473-6783

Journal Of Diabetes & Metabolic Disorders | ISSN: 2381-201X

Journal Of Emergency Medicine Trauma & Surgical Care | ISSN: 2378-8798

Journal Of Environmental Science Current Research | ISSN: 2643-5020

Journal Of Food Science & Nutrition | ISSN: 2470-1076

Journal Of Forensic Legal & Investigative Sciences | ISSN: 2473-733X

Journal Of Gastroenterology & Hepatology Research | ISSN: 2574-2566

Journal Of Genetics & Genomic Sciences | ISSN: 2574-2485

Journal Of Gerontology & Geriatric Medicine | ISSN: 2381-8662

Journal Of Hematology Blood Transfusion & Disorders | ISSN: 2572-2999

Journal Of Hospice & Palliative Medical Care

Journal Of Human Endocrinology | ISSN: 2572-9640

Journal Of Infectious & Non Infectious Diseases | ISSN: 2381-8654

Journal Of Internal Medicine & Primary Healthcare | ISSN: 2574-2493

Journal Of Light & Laser Current Trends

Journal Of Medicine Study & Research | ISSN: 2639-5657

Journal Of Modern Chemical Sciences

Journal Of Nanotechnology Nanomedicine & Nanobiotechnology | ISSN: 2381-2044

Journal Of Neonatology & Clinical Pediatrics | ISSN: 2378-878X

Journal Of Nephrology & Renal Therapy | ISSN: 2473-7313

Journal Of Non Invasive Vascular Investigation | ISSN: 2572-7400

Journal Of Nuclear Medicine Radiology & Radiation Therapy | ISSN: 2572-7419

Journal Of Obesity & Weight Loss | ISSN: 2473-7372

Journal Of Ophthalmology & Clinical Research | ISSN: 2378-8887

Journal Of Orthopedic Research & Physiotherapy | ISSN: 2381-2052

Journal Of Otolaryngology Head & Neck Surgery | ISSN: 2573-010X

Journal Of Pathology Clinical & Medical Research

Journal Of Pharmacology Pharmaceutics & Pharmacovigilance | ISSN: 2639-5649

Journal Of Physical Medicine Rehabilitation & Disabilities | ISSN: 2381-8670

Journal Of Plant Science Current Research | ISSN: 2639-3743

Journal Of Practical & Professional Nursing | ISSN: 2639-5681

Journal Of Protein Research & Bioinformatics

Journal Of Psychiatry Depression & Anxiety | ISSN: 2573-0150

Journal Of Pulmonary Medicine & Respiratory Research | ISSN: 2573-0177

Journal Of Reproductive Medicine Gynaecology & Obstetrics | ISSN: 2574-2574

Journal Of Stem Cells Research Development & Therapy | ISSN: 2381-2060

Journal Of Surgery Current Trends & Innovations | ISSN: 2578-7284

Journal Of Toxicology Current Research | ISSN: 2639-3735

Journal Of Translational Science And Research

Journal Of Vaccines Research & Vaccination | ISSN: 2573-0193

Journal Of Virology & Antivirals

Sports Medicine And Injury Care Journal | ISSN: 2689-8829

Trends In Anatomy & Physiology | ISSN: 2640-7752

Submit Your Manuscript: https://www.heraldopenaccess.us/submit-manuscript

Herald Scholarly Open Access, 2561 Cornelia Rd, #205, Herndon, VA 20171, USA.
Tel: +1 202-499-9679; E-mail: info@heraldsopenaccess.us
http://www.heraldopenaccess.us/