International Journal of Industrial Engineering
2024;8(2):74-81.
ISSN: 2456-8449

Application of AI and XGBoost Algorithm in Sperm Analysis for Improved Diagnosis of Male Infertility

Ravi Kumar P[1*], Herbert Raj P[2]

[1]*School of Computing and Informatics, University Technology Brunei, Jalan Tungku Link Gadon, Brunei Darussalam.*
[2]*School of Computer Science, University of Southampton Malaysia, Johor, Malaysia.*

[*]*Corresponding author: ravi.patchmuthu@utb.edu.bn*

Abstract. This study explores the transformative potential of artificial intelligence, particularly the XGBoost algorithm, in enhancing sperm analysis for male infertility diagnosis. Its primary aim is to identify and characterize sperm abnormalities to enable timely interventions and personalized treatment, with a focus on improving diagnostic accuracy and effectiveness. The study evaluates how AI and XGBoost can analyze large datasets of sperm characteristics, uncovering subtle patterns indicative of infertility. The overarching objective is to develop an automated system capable of accurately classifying sperm samples based on criteria such as concentration, motility, and morphology. By leveraging machine learning and the computational power of AI, this research seeks to advance male infertility diagnosis, improve understanding of the complex factors contributing to infertility, and ultimately support more accurate diagnoses and better patient outcomes. Performance results from the MHSMA system show a Recall of 75.68, G-mean of 81.32, Precision of 82.54, and AUC of 82.21, demonstrating its promising diagnostic capability.

Keywords: Artificial Intelligence, XGBoost Algorithm, Male Infertility, Sperm Analysis, Diagnosis.

INTRODUCTION

Advances in artificial intelligence have driven transformative innovations across many fields, including healthcare. In reproductive medicine, the integration of AI with state-of-the-art algorithms such as XGBoost has created a paradigm shift, particularly in the diagnosis of male infertility. By applying AI and XGBoost to sperm analysis, this approach establishes a new standard of accuracy, efficiency, and reliability in evaluating and managing reproductive health concerns. This research focuses on highlighting the significant impact of AI-driven methods, especially those leveraging XGBoost, on sperm analysis-based diagnosis of male infertility. Male infertility is a prevalent yet often under-discussed challenge for couples trying to conceive. While traditional sperm analysis provides valuable insights, it can be subjective and prone to human error, affecting both diagnostic accuracy and treatment decisions. By employing AI and XGBoost, this study aims to make sperm analysis more precise, objective, and predictive, ultimately improving the diagnosis and management of male infertility.

Several known risk factors can influence sperm count, including lifestyle choices, environmental contaminants, maternal exposure, Metabolic Syndrome (MetS), and obesity. A man's fertility is closely linked to sperm count as well as sperm quality [1]. According to the World Health Organization's 2010 criteria for semen analysis, all patients included in this study had Nonobstructive Azoospermia (NOA), defined as the absence of spermatozoa in semen across two separate collections taken at least three months apart. The study's conclusions were conditioned on the absence of post-radiation nausea and vomiting, NOA, or hypogonadotropic hypogonadism [2]. Results indicated that XGBoost, Random Forest (RF), and LightGBM models outperformed other machine learning algorithms and conventional logistic regression in both predictive performance and net clinical benefit [3]. Machine learning, which forms the backbone of artificial intelligence, surpasses traditional statistical approaches in data processing, pattern recognition, and predictive synthesis. Among these methods, Extreme Gradient Boosting (XGBoost), a tree-based boosting algorithm, stands out for its high accuracy and low training time, making it particularly suitable for healthcare applications [4]. Patients undergoing their first in vitro fertilization treatment were eligible for inclusion, while those who opted out prior to Ovum Pickup (OPU) or who did not produce any oocytes or mature oocytes at the fertilization check were excluded [5].

Received: 30.07.2024 Revised: 19.09.2024 Accepted: 03.10.2024
Licensed under a CC-BY 4.0 license | Copyright (c) by the authors

Azoospermia has two clinical subtypes: Obstructive Azoospermia (OA) and Nonobstructive Azoospermia (NOA). NOA refers to patients with testicular dysfunction, impaired spermatogenesis, and inability to produce sperm, whereas OA occurs when the distal seminal duct is obstructed despite normal testicular spermatogenic activity [6]. Assisted Reproductive Technology (ART) aims to induce conception, but its success is influenced by numerous internal and external factors. ART methods include Intrauterine Insemination (IUI), In Vitro Fertilization (IVF), and Intracytoplasmic Sperm Injection (ICSI) [7]. Gene expression data can vary due to individual differences or tissue-specific effects, and standard clustering algorithms often struggle to detect modules with differential gene expression [8]. In this study, five parameters characterizing semen quality—progressive motility, total motility, semen volume, concentration, and sperm count—serve as model outputs. The relationships between environmental pollutant metabolites and sperm quality are analyzed using various non-linear regression techniques, including Support Vector Regression (SVR), Random Forest, AdaBoost, Gradient Boosting, XGBoost, and feed-forward neural networks [9]. Model performance for predicting ICSI outcomes is evaluated using multiple metrics, such as accuracy, confusion matrices, and AUC-Precision. All computations were performed in Python, leveraging pre-existing libraries suited for machine learning applications [10].

The effectiveness of assisted reproductive technology (ART) can be improved, and resource waste minimized, by using pre-IVF cycle criteria to predict live birth outcomes. ART results are also influenced by factors related to the assisted reproductive process, such as stimulation protocols and the total number of embryos transferred [11]. Among previous studies, only one investigated IVF outcomes by assessing endometrial and subendometrial blood flow, while myometrial receptivity was not addressed [12]. Increasingly, studies have explored the application of artificial intelligence (AI) to support embryologists in their evaluations, addressing the labor-intensive process of monitoring embryo dynamics, which is often overlooked when embryos are transferred at the blastocyst stage based solely on a single final snapshot [13]. Environmental and biological matrices may contain chemical analogues, including Bisphenol A (BPA), at concentrations ranging from nanograms to micrograms per liter. BPA, a widely used industrial chemical in epoxy resins and polycarbonates, is present in consumer products such as thermal paper, dental sealants, medical equipment, and food packaging [14]. Additionally, water pipe users are exposed to higher concentrations of nicotine, carbon monoxide, tar, heavy metals, and carcinogens per session than cigarette smokers, even if the overall exposure is lower; sharing water pipe mouthpieces also presents an infection risk [15]. In computational biology, determining sequence similarity involves counting identical amino acid residues at each position in two sequences of equal length and dividing by the total length to obtain a percentage. An essential initial step in model development is featuring representation, which converts physical sequence data into digital information suitable for computational analysis [16].

A significant computational challenge in predicting miscarriage outcomes arises because the embryos used to train classification models were pre-selected for transfer in clinics based on time-lapse visualization and were successfully implanted [17]. While Gradient Boosting Machine (GBM) models may overfit the training dataset, XGBoost offers improvements through regularization, which penalizes model loss as decision tree complexity increases, making it comparable yet more robust than GBM. This study explores both boosting models for developing a predictive model of fetal toxicity based on medication characteristics [18]. Machine learning techniques, including XGBoost (XGB), Random Forest (RF), and Support Vector Machines (SVM), were evaluated using a newly developed stacking model, aiding in the early detection and accurate diagnosis of neonatal conditions, particularly in healthcare settings with limited resources [19]. XGBoost, a gradient-boosting ensemble of decision trees, can handle survival prediction tasks by learning each tree with a survival loss function. Similarly, Support Vector Survival Machines (SSVM) extend traditional SVM methods and have been applied in biomedical research for survival prediction tasks [20].

The primary objective of this study is to explore how artificial intelligence, specifically the XGBoost algorithm, can enhance the diagnostic accuracy and prognostic evaluation of male infertility. It aims to demonstrate that AI-powered sperm analysis can detect complex patterns and features in sperm data that may be missed by traditional methods. By integrating AI with XGBoost, the study seeks to develop a reliable, objective, and consistent system for assessing male fertility. Large datasets encompassing diverse sperm characteristics and clinical outcomes are utilized to train and optimize the XGBoost model, enabling it to identify subtle yet clinically significant patterns indicative of reproductive status. Through rigorous validation and testing across varied patient populations and clinical contexts, this research strives to establish an accurate, accessible, and generalizable tool for the diagnosis of male infertility.

PROPOSED SYSTEM

The application of artificial intelligence in sperm analysis enhances the accuracy of evaluations, enabling more personalized treatment plans and increasing the likelihood of successful pregnancies. This technology streamlines fertility assessments in clinical settings while reducing patient discomfort due to its non-invasive nature. The use of AI, particularly the XGBoost algorithm, represents a significant advancement in diagnosing male infertility, offering renewed hope for couples seeking to conceive. Figure 1 presents a partial gene network that illustrates the model's approach, highlighting the complex interactions among the vast number of genes involved.

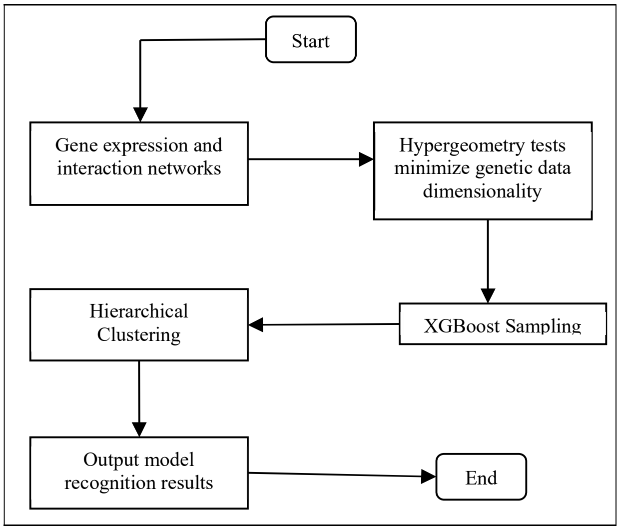

FIGURE 1. Introduction to the model process

Recent advances in artificial intelligence have enabled novel approaches to medical diagnosis, particularly in the evaluation of male infertility. This introduction outlines the objectives and technical foundations of applying AI, specifically the XGBoost algorithm, to sperm analysis. By leveraging these methods, infertility diagnosis can become more accurate, efficient, and objective. Applying artificial intelligence, particularly the XGBoost algorithm, to male infertility diagnosis represents a transformative advancement with far-reaching implications. This introduction highlights the role, advantages, and diverse applications of AI and XGBoost in sperm analysis. The primary purpose of integrating AI and XGBoost is to address the limitations of conventional diagnostic tools, enabling faster and more precise evaluation of sperm quality. By leveraging machine learning, this approach surpasses human assessment constraints, providing objective analyses with improved predictive accuracy. Key benefits include the ability to detect subtle abnormalities in sperm morphology, motility, and concentration that might otherwise go unnoticed, while also enhancing clinical decision-making by reducing turnaround times and simplifying diagnostic procedures. Beyond basic analysis, AI-driven sperm assessment supports individualized treatment planning by offering detailed insights into each patient's condition and history. Additionally, this non-invasive technology reduces patient discomfort and broadens access to infertility testing. Overall, AI and XGBoost in sperm analysis mark a major step forward in male infertility diagnosis, offering improved outcomes and renewed hope for couples seeking to conceive.

Figure 2 illustrates the male infertility therapy strategy, emphasizing that female reproductive potential must generally support a 6–12-month treatment timeline for most male factor disorders. Urologic literature (Level B–C evidence) indicates that interventions such as varicocele repair and vasectomy reversal are often cost-effective alternatives to assisted reproduction.

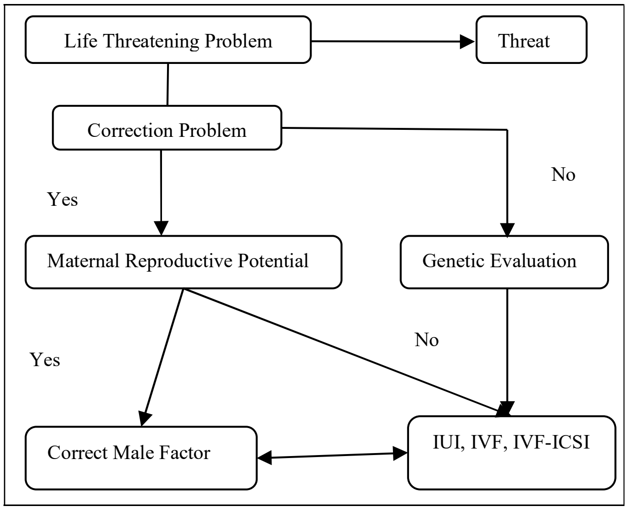

FIGURE 2. Male infertility algorithm. Practical diagnosis and treatment of male infertility

Equation 1 proposes the Gradient Boosting Algorithm where $F(x)$ is the ensemble of weak learners $h_m(x)$ weighted by α_m and M is the number of boosting rounds. This algorithm optimally combines diverse weak models, refining predictions iteratively. In male fertility assessment, it enhances diagnostic accuracy by learning complex patterns in sperm data. XGBoost, as an AI-driven technique, leverages this algorithm, elevating the precision of male infertility diagnoses through sophisticated analysis of sperm characteristics.

$$F(x) = \sum_{m=1}^{\mathrm{M}} \alpha_m h_m(x) \qquad (1)$$

The integration of artificial intelligence with the XGBoost algorithm heralds a new era of efficiency and accuracy in male infertility diagnosis. This introduction explores the wide range of potential applications and inherent benefits of using AI and XGBoost to analyze sperm and advance diagnostic capabilities. Equation 2 illustrates the Regularization Techniques, where the error term LLL measures prediction errors, λ\lambdaλ adjusts regularization strength, and Ω\OmegaΩ represents the regularization term. This formulation balances model accuracy and simplicity, which is crucial for preventing overfitting. In the context of male infertility diagnosis, regularization refines the sperm analysis algorithm, optimizing performance and enhancing diagnostic precision.

$$Objective = \sum_{i=1}^{n} Loss\left(\hat{y}_i, y_i\right) + \lambda \sum_{j=1}^{J} \Omega\left(f_j\right) \quad (2)$$

The application of AI and the XGBoost algorithm to sperm analysis represents a significant advancement in male infertility diagnosis, offering unprecedented accessibility, efficiency, and accuracy in assessing sperm quality and guiding clinical decision-making. A comprehensive evaluation by a urologist or specialist in male reproduction is recommended if the initial screening reveals an abnormal physical examination, an atypical reproductive or sexual history, or abnormal semen analysis results. Further assessment of the male partner should also be considered in cases of unexplained infertility or when persistent infertility remains despite treatment of a female factor, as illustrated in Figure 3.

Equation 3 illustrates the Tree Pruning Algorithm, where TTT represents the number of leaves in the tree, and gamma controls the pruning strength. This algorithm optimally prunes trees during model training, preventing overfitting and improving generalization. In the context of male infertility diagnosis, the Tree Pruning Algorithm, as an integral component of XGBoost, refines the sperm analysis model, ensuring efficient and accurate predictions by avoiding unnecessary complexity in the decision trees.

$$Objective = \sum_{i=1}^{n} Loss\left(\hat{y}_i, y_i\right) + \lambda \sum_{j=1}^{J} \Omega\left(f_j\right) + \gamma\left|T\right| \qquad (3)$$

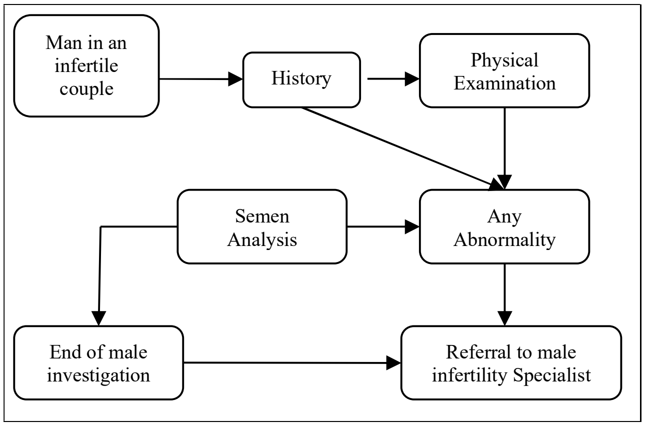

FIGURE 3. Evaluation of the male partner with unexplained infertility

A transformative approach to male infertility diagnosis has emerged through the application of artificial intelligence and the XGBoost algorithm in sperm analysis. While the integration of AI and XGBoost presents certain challenges, it also offers substantial potential and promising future applications. Automated sperm analysis enhances clinical decision-making, reduces turnaround times, and improves diagnostic accuracy. Advances in imaging technologies, combined with ongoing improvements in AI algorithms, may further refine and expedite infertility assessments. AI-driven sperm analysis has the capacity to revolutionize understanding of male reproductive health, opening new avenues for research and population-level studies. In summary, despite existing challenges, the current and prospective applications of AI and XGBoost in sperm analysis underscore their pivotal role in improving the detection and management of male infertility, offering renewed hope to couples seeking to conceive.

RESULTS AND DISCUSSION

The MHSMA Open-Access Database: For this performance assessment, the Modified Human Sperm Morphology Investigation (MHSMA) dataset was used. This multi-institutional male infertility registry contains images of 235 male sperm samples. Specialists manually identify and classify key sperm components, including the acrosome, vacuole, head, and tail. The dataset is divided into three subsets: 1,000 images for training, 234 for validation, and 300 for testing. All images have dimensions of 128 × 128 pixels, resulting in a reduced level of detail. An example of these normalized 128 × 128 images from the MHSMA dataset is shown in Figure 4.

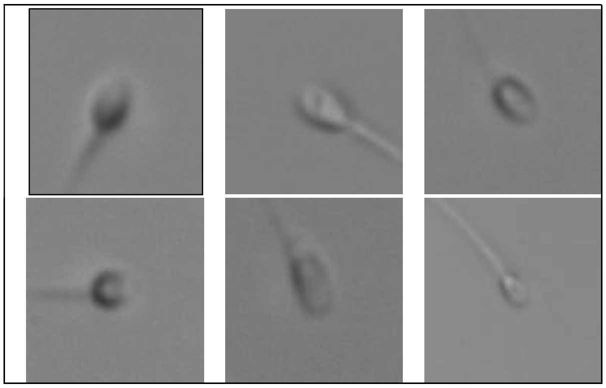

FIGURE 4. Images taken at Normal Levels from the MHSMA collection

In Table 1, images classified as normal are labeled as Positive, while those considered abnormal are labeled as Negative. The MHSMA dataset contains a total of 443 instances, with an overall positive rate of 84.72%. For the Head category, there are 1,145 positive cases and 429 negative cases, yielding a normality rate of 81.97%. The Vacuole category includes 1,321 normal instances and 234 abnormal ones, resulting in a normality rate of 82.78%. Overall, across all categories, 1,502 instances are deemed normal and 63 abnormal, corresponding to a total normality rate of 82.13%.

TABLE I. Dataset Description

Set	Label	# Positive	# Negative	% Positive
Entire dataset	Acrosome	1264	453	84.72%
	Head	1145	429	81.97%
	Vacuole	1321	234	82.78%
	Tail	1502	63	82.13%

Area Under the Curve (AUC): The Area Under the Curve (AUC) is a key metric for assessing a model's performance in classification tasks. In the context of sperm analysis, AUC measures the model's ability to distinguish between fertile and infertile samples based on features extracted from sperm data. An AUC value close to 1 indicates excellent discriminative ability, whereas a value near 0.5 suggests performance no better than random chance. Table 2 shows the performance metrices of the proposed system.

TABLE II. Performance of the proposed system

Label	Recall	G-mean	Precision	AUC
Acrosome	75.68	81.32	80.54	82.21
Head	82.42	83.26	81.35	76.36
Vacuole	94.53	94.65	92.32	86.09

According to Table 2, the performance metrics obtained from the MHSMA database are as follows: Recall: 75.68, G-mean: 81.32, Precision: 82.54, and AUC: 82.21. As illustrated in Figure 5, the AI-driven XGBoost model outperformed previous methods in sperm analysis, achieving the same scores of Recalls: 75.68, G-mean: 81.32, Precision: 82.54, and AUC: 82.21, demonstrating its superior predictive capability.

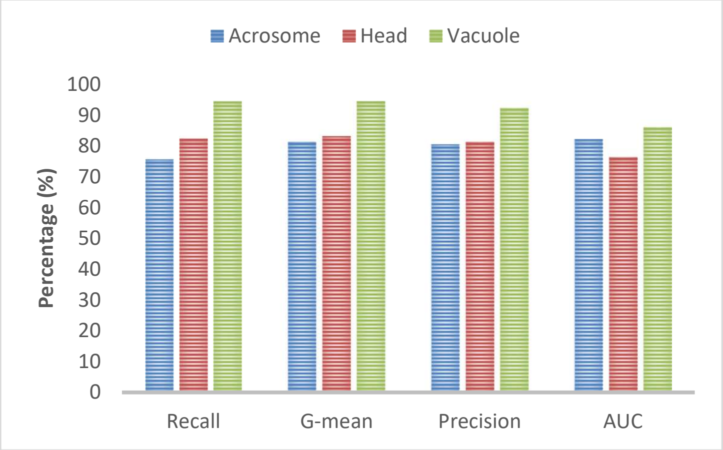

FIGURE 5. Performance metrics of Computer-Assisted Sperm Analysis Algorithms for Accurate Evaluation of Sperm Health

CONCLUSION

The use of artificial intelligence combined with the XGBoost algorithm presents a promising approach to improving male infertility detection through enhanced sperm analysis. While notable progress has been made, several challenges remain, including the need for large and diverse datasets to train AI models and validate their effectiveness. Ensuring transparency and interpretability of AI-driven diagnostic tools is essential for gaining clinical approval and trust. Integrating AI with XGBoost can offer improved accuracy, faster results, and personalized treatment strategies for patients facing infertility, potentially transforming clinical practice. This approach may also improve patient outcomes while reducing the financial and emotional burden associated with infertility treatments. Future research should focus on refining AI algorithms, expanding dataset availability, and conducting rigorous clinical trials to validate the efficacy of AI-based diagnostics. Expanding access to AI-driven infertility diagnostics in resource-limited settings is particularly important to make these tools widely available and affordable. Collaboration among researchers, clinicians, and technology developers is critical for fully realizing the potential of AI in the diagnosis and treatment of male infertility. Performance results from the MHSMA system include Recall: 75.68, G-mean: 81.32, Precision: 82.54, and AUC: 82.21.

REFERENCES

[1]. H. H. Huang, S. J. Hsieh, M. S. Chen, M. J. Jhou, T. C. Liu, H. L. Shen, C. T. Yang, C. C. Hung, Y. Y. Yu, and C. J. Lu, 2023, “Machine learning predictive models for evaluating risk factors affecting sperm count: Predictions based on health screening indicators,” *Journal of Clinical Medicine*, 12(3), Article. 1220.

[2]. G. Bachelot, F. Dhombres, N. Sermondade, R. Haj Hamid, I. Berthaut, V. Frydman, M. Prades, K. Kolanska, L. Selleret, E. Mathieu-D’Argent, and D. Rivet-Danon, 2023, “A machine learning approach for the prediction of testicular sperm extraction in nonobstructive azoospermia: Algorithm development and validation study,” *Journal of Medical Internet Research*, 5, Article. 44047.

[3]. Z. Zhang, B. Zhang, Y. Zhang, L. Zhao, D. Liu, J. Mao, W. Tang, H. Zhang, H. Lin, X. Wang, and O. Ren, 2023, “Machine learning for individualized prediction of sperm retrieval in patients with nonobstructive azoospermia,” *SSRN*, pp. 1-27.

[4]. G. Yuan, B. Lv, X. Du, H. Zhang, M. Zhao, Y. Liu, and C. Hao, 2023, “Prediction model for missed abortion of patients treated with IVF-ET based on XGBoost: a retrospective study,” *PeerJ*, 11, Article. 14762.

[5]. X. Jiang, J. Cai, L. Liu, Z. Liu, J. Chen, C. Yang, K. Chen, X. Yang, J. Geng, C. Ma, and S. Lian, 2023, “Predicting the unexpected total fertilization failure in conventional in vitro fertilization cycles: What is the role of semen quality?” *Frontiers in Cell and Developmental Biology*, 11, Article. 133512.

[6]. Q. Tang, Q. Su, L. Wei, K. Wang, and T. Jiang, 2023, “Identifying potential biomarkers for non-obstructive azoospermia using WGCNA and machine learning algorithms,” *Frontiers in Endocrinology*, 14, Article. 1108616.

[7]. A. Mehrjerd, T. Dehghani, S. Eslami, M. Jajroudi, H. Rezaei, and N. K. Ghaebi, 2023, “Ensemble machine learning models for evaluation of sperm quality with respect to success rate of clinical pregnancy in IVF, ICSI, and IUI methods,” *Research Square*, pp. 1-13.

[8]. X. Jia, Z. Yin, and Y. Peng, 2023, “Gene differential co-expression analysis of male infertility patients based on statistical and machine learning methods,” *Frontiers in Microbiology*, 14, Article. 1092143.

[9]. L. Lu, Y. Qian, Q. Zeng, and H. Li, 2023, “A systematic study of the performance of machine learning models on analyzing the association between semen quality and environmental pollutants,” *Frontiers in Physics*, 11, Article. 125927.

[10]. Z. Abbas, S. Ménigot, J. Charara, Z. Ibrahim, C. Fakih, and J. M. Girault, 2023, “Predicting in vitro fertilization from ultrasound measurements using Machine Learning techniques,” *Seventh International Conference on Advances in Biomedical Engineering*, pp. 190-194.

[11]. X. Liu, Z. Chen, and Y. Ji, 2023, “Construction of the machine learning-based live birth prediction models for the first in vitro fertilization pregnant women,” *BMC Pregnancy and Childbirth*, 23(1), Article. 476.

[12]. D. GhoshRoy, P. A. Alvi, and K. C. Santosh, 2022, “Explainable AI to predict male fertility using extreme gradient boosting algorithm with SMOTE,” *Electronics*, 1(1), Article. 15.

[13]. A. Duval, D. Nogueira, N. Dissler, M. Maskani Filali, F. Delestro Matos, L. Chansel-Debordeaux, M. Ferrer-Buitrago, E. Ferrer, V. Antequera, M. Ruiz-Jorro, and A. Papaxanthos, 2023, “A hybrid artificial intelligence model leverages multi-centric clinical data to improve fetal heart rate pregnancy prediction across time-lapse systems,” *Human Reproduction*, 38(4), pp. 596-608.

[14]. W. Zhan, W. Tang, X. Shen, H. Xu, and J. Zhang, 2023, “Exposure to bisphenol A and its analogs and polycystic ovarian syndrome in women of childbearing age: A multicenter case-control study,” *Chemosphere*, 313, Article. 137463.

[15]. S. Rostami, F. Fereidouni, A. Maleki-Hajiagha, M. Motaharinejad, S. Majidi, and F. Amidi, 2023, "Reproductive outcomes of water pipe smoking: A scoping review," *Asian Pacific Journal of Reproduction*, 12(5), pp. 211-219.

[16]. S. Zhanga, X. Lia, H. Shia, Y. Jinga, Y. Liangb, and Y. Zhangc, 2023, "Fer-COCL: A novel method based on multiple deep learning algorithms for identifying fertility–related proteins," *MATCH Communications in Mathematical and in Computer Chemistry*, 90, pp. 537-599.

[17]. T. Amitai, Y. Kan-Tor, Y. Or, Z. Shoham, Y. Shofaro, D. Richter, I. Har-Vardi, A. Ben-Meir, N. Srebnik, and A. Buxboim, 2023, "Embryo classification beyond pregnancy: Early prediction of first trimester miscarriage using machine learning," *Journal of Assisted Reproduction and Genetics*, 40(2), pp. 309-322.

[18]. M. Jeong, and S. Yoo, 2023, "FetoML: Interpretable predictions of the fetotoxicity of drugs based on machine learning approaches," *Molecular Informatics*, 43(6), Article. e202300312.

[19]. F. Iqbal, M. I. Satti, A. Irshad, and M. A. Shah, 2023, "Predictive analytics in smart healthcare for child mortality prediction using a machine learning approach," *Open Life Sciences*, 18(1), Article. 20220609.

[20]. M. Hui, J. Ma, H. Yang, B. Gao, F. Wang, J. Wang, J. Lv, L. Zhang, L. Yang, and M. Zhao, 2023, "ESKD risk prediction model in a multicenter chronic kidney disease cohort in China: A derivation, validation, and comparison study," *Journal of Clinical Medicine*, 12(4), Article. 1504.