Preliminary experience on the use of artificial intelligence-assisted sperm search in surgical sperm retrieval.

Saud Ibrahim Sharif Hussain[1], and Khalid AlMuqati[1]

[1] IVF Laboratory, Dr Sulaiman AlHabib Hospital As Suwaidi, Riyadh, Kingdom of Saudi Arabia

Abstract

The objective was to validate the SpermSearchAI™ software. Observation 1: The protocol design was identical to manufacturer's. Observation 2: The dish was flooded with a thin layer of medium containing a mixture of spermatozoa and testicular tissue cells. Observation 1 attempted to detect a clean specimen of extremely low concentrations of washed neat spermatozoa (NS) held in a HEPES-buffered medium devoid of any cells. Subsequent investigations repeated the observation with TESA (TE) specimen. Observation 1 noted the AI detected spermatozoa faster, but the human eye detected significantly more ($p=0.0085$) spermatozoa, mean spermatozoa counted (MSCs):13.8 vs 8.4 respectively, n=20, correlation coefficient $r=0.9285$, $p<0.0001$, while the mean false positive observations (MFOs) were significantly higher in AI-group, mean 0.2 vs 0.0, $p<0.05$, with no MFOs in the human. In TE specimen, the MSCs were similar in human and AI observations: 7.9 vs 8.0, $p=0.7054$, respectively, n=20, $r=0.9558$, $p<0.0001$; MFOs were similar to the previous observation, mean 3.0 vs 0.0, $p<0.0001$. TE specimen without dilution, mean spermatozoa counts were 3.35 vs 1.2, $p<0.0001$, for human and AI respectively, n=20, $r=0.6485$, $p=0.0020$, the MPOs were similar to previous observations, mean 2.3 vs 0.0, $p<0.0001$); so were results of corresponding diluted TE specimen, the difference between mean spermatozoa counts were not significant, 0.78 vs 0.70, $p=0.1619$, n=23, $r=0.9221$, $p<0.0001$, the MPOs were similar to previous observations mean 1.3 vs 0.0, $p<0.0001$. When the observations of all observations, except the first were combined, the mean of spermatozoa counts were significantly higher ($p=0.0003$) in the human compared to AI respectively: 3.9 vs 3.2, n=63, $r=0.9353$, $p<0.0001$, while the MPOs were similar to previous observations, mean 2.2 vs 0.0, $p<0.0001$. In conclusion, AI detects spermatozoa faster but with significantly high false positive observations than the human but the human counts significantly more spermatozoa with no false positive observations.

Disclaimer: The authors declare no conflict of interest.
J Reprod Biotechnol Biomed Sci.15:1-6
Correspondence: Saud Ibrahim Sharif Hussain**; Email:** www.saud.899@hotmail.com
Compliance acknowledgement: This article was edited by the Australian Editorial Services (www.nativeenglisheditor.com)
Keywords: Artificial intelligence, assisted, search, spermatozoa, testicular
Notification: The statements and claims in this manuscript are that of the author(s) alone.

Introduction

It is well documented the male is responsible for around 30-50% of infertility. For many couples with male factor infertility the use of intracytoplasmic sperm injection (ICSI) remains the most common technique employed to achieve fertilization (Eisenberg et al., 2023). We now know male infertility is increasing at an alarming rate and in fact has dropped by 50% in the last 50 years (Levine et al., 2023). In severe oligozoospermia and cryptozoospermia the standard and well documented method of sperm preparation, recovery and its selection for insemination by ICSI are by density gradient centrifugation, followed by manual sperm detection under inverted microscope and finally selection of morphologically normal sperm for insemination respectively.

Non-obstructive azoospermia (NOA) affects about 60% of azoospermic men (Jarow et al., 1989; Wosnitzer, et al., 2014). The current ART treatment for patients with NOA is ICSI with

sperm extracted from the testis. The laboratory staff are confronted with a specimen containing many blood cells, testicular cells and tissues, and a small number of testicular sperm (TS). Oftentimes the number of sperm in such specimen may be extremely few in NOA. In obstructive azoospermia the number of sperm recovered is a little higher (unpublished observations). To recover and select sperm from testicular specimen with extremely limited numbers of sperm poses challenges to the ICSI operator.

It is not uncommon for workers to spend as long as 2 hours or more to recover a few sperm for ICSI insemination (Mangum et al., 2020; Ramasamy et al., 2011). It is well recognised among the fraternity the long hours needed to recover sperm from TS could result in work-related stress or failure to recover sperm due to fatigue-induced human error. In extreme situation it is also known to cause severe burn-out in some ICSI operators, especially in busy under-staffed laboratories.

Fatigue-induced human error resulting in failure to detect sperm in NOA patients would lead to the wrong diagnosis of absolute infertility (Samuel et al., 2016), oftentimes with devastating emotional sequelae for the affected patient. Equally harmful situations are known to occur in patients with severe oligozoospermia and cryptozoospermia. Failure to detect sperm in these men, or the longer time taken for sperm searches for diagnosis or in instances of a final examination before surgery, may lead to unnecessary surgery. In the ICSI procedure, oftentimes prolonged sperm search could affect the fertilizing potential of the sperm. These inadequate attempts in the laboratory caused by human error undermine ART treatment procedures at enormous unproductive costs and, physical and financial strain to the affected patients (Ouitra-kul et al., 2018), often accompanied by considerable emotional pressure.

With the advent of artificial intelligence (AI) it is inevitable that it will be utilized in various areas of healthcare ((Elias et al., 2023), one of which is AI-assisted sperm search. Proponents of AI-powered image analysis claim it has the potential to improve ART treatment (McCallum et al., 2019; Wang et al., 2019; Mendizabal-Ruiz et al., 2022; Joshi et al., 2023) and reduce significantly the time to detect and recover spermatozoa from TS in ART (Goss et al., 2023, 2024; Ricarte et al., 2026).

We describe here our preliminary experience with the SpermSearchAI™ software (NeoGenix Fertility Pty Ltd, Australia) for the detection of sperm in TS. The present finding is that of validation data obtained during routine TS detection.

Methodology

The question was what the difference between the human eye and AI is when presented with a neat and clean specimen of washed spermatozoa held in HEPES-buffered holding medium (Cooper Surgical, Origio Cat no. 10760125M) that is not accompanied by potentially confounding blood and testicular cells or testicular tissues (validation observation 1) and when challenged with spermatozoa plus blood and testicular cells, and testicular tissues (validation observation 2). This evaluation procedure is intended to show which of the two, the human eye or AI is superior in detecting the spermatozoa.

.

Observation 1: The observation dish was prepared identical to manufacturer's instructions. The specimen used was washed spermatozoa in medium devoid of any cells or tissues. Observation 1 attempted to detect a clean specimen of extremely low concentrations of washed neat spermatozoa (NS) held in HEPES-buffered medium devoid of cells. The specimen was not challenged with confounding blood and testicular germ cells, and tissues normally found in TS.

The spermatozoa count in the washed specimen was 105 million per ml. A serial dilution performed 14 times with each dilution halving the concentration of the sperm led to a very low count of approximately 0.006 million sperms per ml. This 14^{th} dilution was used for the validation Observation 1 procedure.

For the human group, 2 workers scanned 10 fields each, whereas the AI readings were recorded by the same two workers, 10 fields each, alternating between manual and AI-assisted sperm search performed at the magnification of 40x using a Nikon Eclipse Ti model inverted microscope. This effort resulted in 20 replicated readings per treatment group.

Observation 2: The investigation is repeated with MicroTESA (mTE) and TESA (TE) specimens. The dish was flooded with a thin layer of medium containing a mixture of spermatozoa and autologous testicular tissue cells during routine microTESA (mTE) which had few blood and germ cells and a TESA specimen which had more sperm, blood and testicular cells. In a subsequent observation, the mTE and TE specimen were combined and scanned as previously described. The number of replicates readings were 20 per set of observations. We routinely perform about 970 TE and mTE per year.

The statistical analyses were performed with the statistical software, MedCalc®

Results

The results of the first observation involving washed sperm without blood cells or testicular tissues noted the AI detected spermatozoa faster by about 5 seconds, but the human eye detected significantly more (p=0.0085) spermatozoa counts, mean spermatozoa counted (MSCs) were 13.8 vs 8.4 respectively, n=20, correlation coefficient r=0.9285 (Figs. 1& 2), p<0.0001.

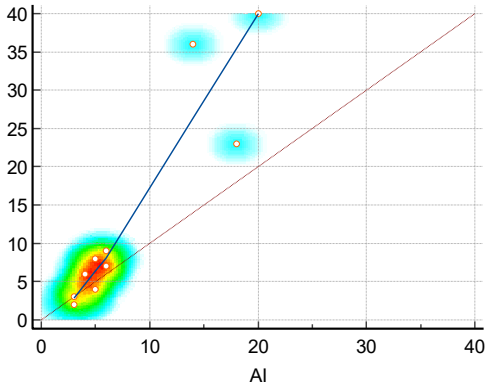

Fig 1: Difference in number of sperms detected by the human eye and AI in washed sperm in medium without cells ro tissues.

Whereas the mean false positive observations (MFOs) were significantly higher in the AI-group, mean 0.2 vs 0.0, p<0.05, with no MFOs in the human.

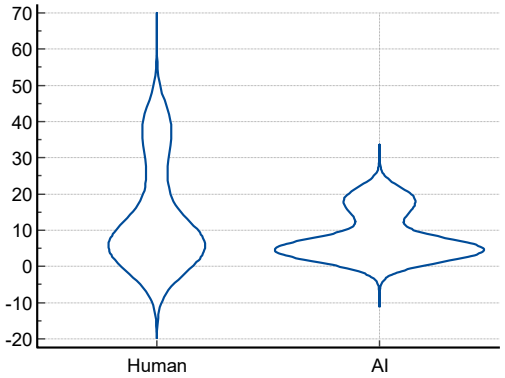

Fig 2: Violin plot between human and AI for the numbers of washed spermatozoa in medium detected.

In the mTE specimen, the MSCs were similar in the human and AI observations: 8.0 vs 7.9, but the difference was insignificant p=0.7054 respectively, n=20, r=0.9558, (figs3&4) p<0.0001, the FPOs were similar to the previous observation, mean 3.0 vs 0.0, p<0.0001.

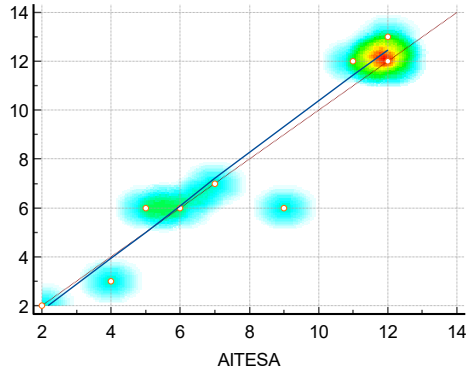

Fig 3: Difference in number of sperms detected by the human eye and AI in diluted testicular specimen

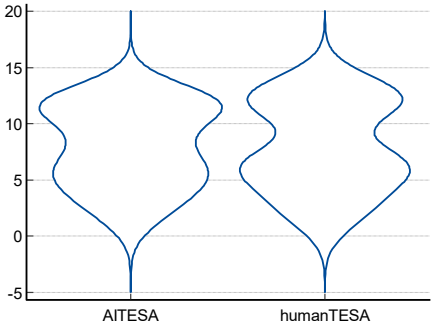

Fig 4: Violin plot of the difference in number of sperms detected by the human eye and AI in diluted testicular specimen

TE specimen without dilution which contained much more blood, testicular cells and tissues, the mean spermatozoa counts were 3.35 vs 1.2, p<0.0001, for the human and AI respectively, n=20, r=0.6485 (Figs.5 & 6).

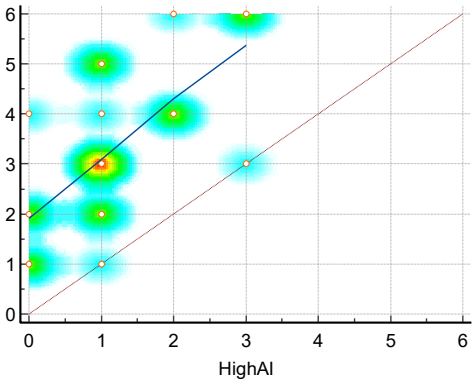

Fig 5: Difference in number of sperms detected by the human eye and AI in undiluted neat testicular specimen

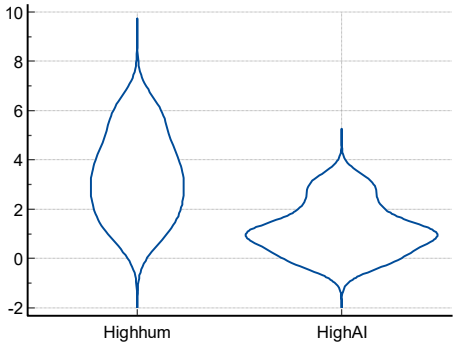

Fig 6: Violin plot of the difference in number of sperms detected by the human eye and AI in undiluted neat testicular specimen

p=0.0020, the FPOs were similar to previous observations, mean 2.3 vs 0.0, p<0.0001); so were the results of the corresponding diluted TE specimen but the difference between mean spermatozoa counts were not significant, 0.78 vs 0.70, p=0.1619, n=23, r=0.9221 (Fig. 1), p<0.0001, the FPOs were similar to the mean of previous observations 1.3 vs 0.0, p<0.0001.

When the mTE and TE specimen, except the washed spermatozoa were combined, the mean of spermatozoa counts were significantly higher (p=0.0003) in the human compared to AI respectively: 3.9 vs 3.2, n=63, r=0.9353, p<0.0001, while the FPOs were similar to previous observations, mean 2.2 vs 0.0, p<0.0001.

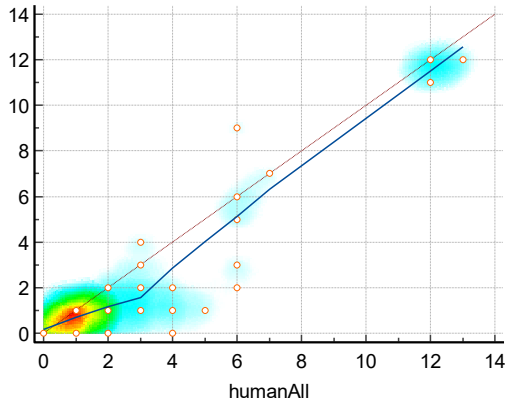

Fig 7: Difference in number of sperms detected by the human eye and AI in all testicular specimen combined

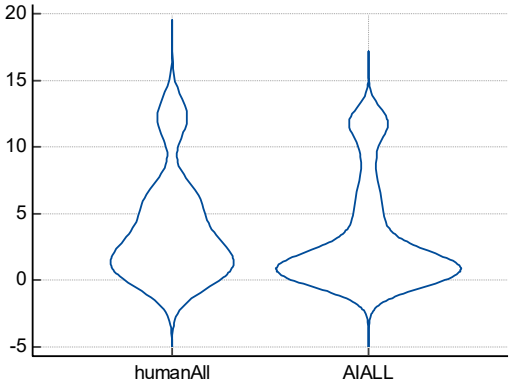

Fig 8a: Violin plot of the difference in number of sperms detected by the human eye and AI in all testicular specimen combined.

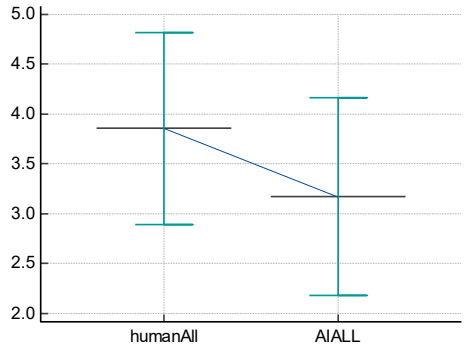

Fig 8b: Comparative plot of the difference in mean and total numbers of sperms detected by the human eye and AI in all testicular specimen combined.

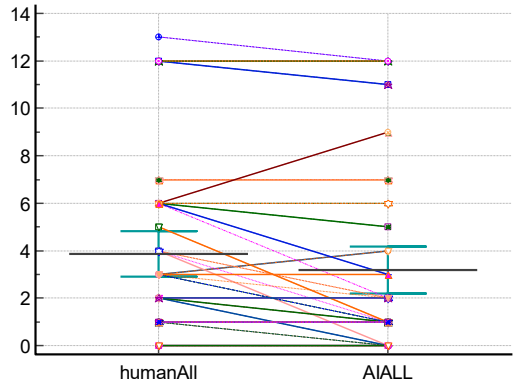

Fig 8c: Stratified plot of the difference in the individual numbers of sperms detected per observation by the human eye and AI in all testicular specimen combined.

Discussion

The findings of this preliminary validation report appear to favor the human over the AI when it comes efficacy in detecting spermatozoa. It is almost certain the human is as good as if not better than the AI in detecting the spermatozoa if the human has had extensive exposure and training in sperm detection in testicular specimen.

It is clear the AI is a little faster by a mere 5 seconds, but the human was able to detect far more spermatozoa with no false positive detections. Although there were no confounding blood and testicular cells or testicular tissue the AI was not superior in its capability to detect sperm in a neat and clean specimen containing only sperm, which further strengthens our assumption that the human eye could better the AI.

However, this is only a preliminary validation report involving a small sample size. We must also point out we are a specialized clinic devoted to male factor infertility catering to patients from not only Saudi Arabia but the entire GCC (Gulf Cooperation Countries) and other parts of the Middle East. We routinely perform about 970 testicular spermatozoa searches per year. Consequently, therefore we have considerable hands-on experience with testicular specimen which probably places us at an advantage. We will be undertaking a much larger observation to confirm the present findings.

The AI-assisted spermatozoa search software could be beneficial in clinics where the workers experience a much lower number of testicular spermatozoa search compared to the workers of this report.

Conclusion

In conclusion, AI detects spermatozoa at an insignificant 5 seconds faster than the human eye but with significantly high false positive observations. The trained human eye counts significantly more spermatozoa with no false positive observations compared to the AI.

Nature of Contributions

Intellectual contributions (92%) were provided by the following: SISH staged and performed the laboratory investigations and, provided inputs in drafting this manuscript which amounted to 48% of the contribution to this validation report. MAK performed the laboratory investigation alongside SISH and assisted in the drafting of this manuscript amounting to 44% of this work.

Technical assistance (8%): was provided by the members of the IVF Laboratory, the nursing personnel, and staff of the Urology Department of this hospital.

Acknowledgement

We would like to express our sincere gratitude to Alaa Gasim Bader Abdullah for her technical assistance in preparations for the validation observation 1.

Reference

Eisenberg ML, Esteves SC, Lamb DJ, Hotaling JM, Giwercman A, Hwang K, Cheng YS. Male infertility. Nat Rev Dis Primers. 2023;9:49. Cited in: Ricarte LL, Cavalcanti RV, Monteiro PB. Use of artificial intelligence in the choice of sperm for ICSI - A literature review. JBRA Assist Reprod. 2026 Apr 1;30(2):396-400

Elias MA, Faversani LA, Moreira JAV, Masiero AV, Cunha NV. Artificial intelligence in health and bioethical implications: a systematic review. Rev Bioét. 2023;31:e3542PT.

Goss D, Vasilescu S, Vasilescu P, Sacks G, Gardner D, Warkiani M. Artificial intelligence to

assist in surgical sperm detection and isolation. Hum. Reprod. 2023;38:163 suppl.

Goss DM, Vasilescu SA, Vasilescu PA, Cooke S, Kim SH, Sacks GP, Gardner DK, Warkiani ME. Evaluation of an artificial intelligence-facilitated sperm detection tool in azoospermic samples for use in ICSI. Reprod Biomed Online. 2024l;49(1):103910.

Jarow, J.P. · Espeland, M.A. · Lipshultz, L.I. Evaluation of the azoospermic patient. J Urol. 1989; 142:62-65

Joshi K, Simbulan RK, Rajah AM. A proof-of-concept prospective observation of applying artificial intelligence for sperm selection in the IVF laboratory. Reprod BioMed Online. 2023;103329.

Levine H, Jørgensen N, Martino-Andrade A. 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. 2023; 29:157-176

Mangum CL, Patel DP, Jafek AR. Towards a better testicular sperm extraction: novel sperm sorting technologies for non-motile sperm extracted by microdissection TESE. Transl Androl Urol. 2020; 9:S206-S214

McCallum C, Riordon J, Wang Y, Kong T, You JB, Sanner S, Lagunov A, Hannam TG, Jarvi K, Sinton D. Deep learning-based selection of human sperm with high DNA integrity. Commun Biol. 2019;2(250):1-10.

Mendizabal-Ruiz G, Chavez-Badiola A, Figueroa IA. Computer software (SiD) assisted real-time single sperm selection associated with fertilization and blastocyst formation. Reprod BioMed Online. 2022; 45:703-711

Ouitrakul S, Sukprasert M, Treetampinich C. The effect of different timing after ejaculation on sperm motility and viability in semen analysis at room temperature. J. Méd Assoc Thail. 2018; 101:26-32

Ramasamy, R. · Reifsnyder, J.E. · Bryson, C. Role of tissue digestion and extensive sperm search after microdissection testicular sperm extraction. Fertil. and Steril. 2011; 96:299-302

Ricarte LL, Cavalcanti RV, Monteiro PB. Use of artificial intelligence in the choice of sperm for ICSI - A literature review. JBRA Assist Reprod. 2026 Apr 1;30(2):396-400

Samuel R, Badamjav O, Murphy KE. Microfluidics: The future of microdissection TESE? Systems Biol Reprod Med. 2016; 62:161-170

Wang Y, Riordon J, Kong T, Xu Y, Nguyen B, Zhong J, You JB, Lagunov A, Hannam TG, Jarvi K, Sinton D. Prediction of DNA integrity from morphological parameters using a single-sperm DNA fragmentation index assay. Adv Sci (Weinh) 2019;6:1900712

Wosnitzer, M. · Goldstein, M. · Hardy, M.P. Review of azoospermia. Spermatogenesis. 2014; 4:e28218