High-throughput imaging cytometry paired with artificial intelligence identifies ultrarare sperm in men previously diagnosed with clinical azoospermia

Blair Stocks, M.D., Ph.D.,[a] Aidan Boyne, B.S.,[b] Amelia Oppenheimer, B.S.,[a] Katelynn Salmon, B.S.,[b] Taylor Kohn, M.D., M.Phil.,[a] Gal Saffati, M.D.,[a] Beatriz Hernandez, M.D.,[a] Justin Badal, M.D.,[a] Mohit Khera, M.D., M.BA., M.P.H.,[a] and Larry Lipshultz, M.D.[a]

[a] Scott Department of Urology, Division of Male Reproductive Medicine and Surgery, Baylor College of Medicine, Houston, Texas; [b] School of Medicine, Baylor College of Medicine, Houston, Texas

Objective: To determine whether a novel diagnostic platform, which pairs high-throughput imaging cytometry with artificial intelligence (AI)–assisted image analysis, can identify ultrarare sperm in ejaculatory samples deemed clinically azoospermic.
Design: A retrospective cohort review of remnant biobanked ejaculatory samples processed by high-throughput imaging cytometry and AI-assisted image analysis.
Subjects: Remnant ejaculatory samples (n = 148) biobanked from December 2024 to September 2025 at a single high-volume male infertility clinic.
Exposure: Image-based flow cytometry and convolutional neural network (CNN) driven image analysis for diagnostic classification.
Main Outcome Measures: Diagnostic reclassification rate of remnant ejaculatory samples initially deemed clinically azoospermic using CNN-based morphologic assessment and corollary sperm specific protein expression.
Results: Within the 83 azoospermic samples analyzed, 40 samples (48%) were found to have sperm using this high-fidelity diagnostic platform.
Conclusion: We propose a new classification of sperm concentration, microzoospermia, a sperm concentration so low that only advanced image-based search technologies can reliably identify ultrarare sperm. (Fertil Steril® 2026;126:28–37. ©2026 by American Society for Reproductive Medicine.)
El resumen está disponible en Español al final del artículo.

Key Words: Imaging cytometry, artificial intelligence, azoospermia

The 2024 American Society for Reproductive Medicine (ASRM)/American Urological Association (AUA) male infertility guidelines recommend upfront microtesticular sperm extraction (mTESE) for men with nonobstructive azoospermia (NOA) undergoing surgical sperm retrieval (1). Reported sperm retrieval rates approach 50% (2–4). Whether one views these chances glass half-full or glass half-empty, there remain two needs. First, a means to predict one's odds of success before undertaking the physical, emotional, and financial burden of this invasive procedure (5). And second, a more reliable and timely means to identify ultrarare sperm in ejaculate and/or mTESE samples.

Efforts to predict mTESE success are ongoing. At the deoxyribonucleic acid (DNA) level, researchers are harnessing long-read genome sequencing to detect the presence of sperm-derived, cell-free DNA in azoospermic ejaculate supernatants (6). At the ribonucleic acid (RNA) level, unique gene expression patterns have correlated with different stages of spermatogenesis matching testicular biopsy pathologies (7). At the protein level, the expression of AKAP4, ACRV1, ZPBP2, PGK2, and TEX101 in azoospermic ejaculates has predicted

Received January 22, 2026; revised April 15, 2026; accepted April 16, 2026; published online April 24, 2026.
B.S. and A.B. should be considered similar in author order.
Supported by Internal funding and the Cytometry and Cell Sorting Core at Baylor College of Medicine with funding from the CPRIT Core Facility Support Award (CPRIT-RP180672), the NIH (P30 CA125123 and S10 RR024574) and the expert assistance of Joel M. Sederstrom, Claude Chew, Mason Perry, and Amy Haggerty.
Correspondence: Larry Lipshultz, M.D. (E-mail: larryl@bcm.edu).

Fertil Steril® Vol. 126, No. 1, July 2026 0015-0282/$36.00

https://doi.org/10.1016/j.fertnstert.2026.04.011

successful surgical sperm retrieval (8–10). Although these emerging noninvasive assays exploring the genomics, transcriptomics, and proteomics of human spermatogenesis hold promise, they have yet to reliably enter the clinical arena. These preliminary studies still mandate further scientific validation and peer review. Accordingly, reproductive urologists rely primarily on clinical nomograms to predict mTESE success (11).

After an mTESE is performed, trained embryologists must search for rare sperm. In many cases, they are unable to find sperm and couples must choose to use donor sperm or freeze unfertilized ova. To address these low-yield results, a single US andrology laboratory offers extended sperm search ± microfreeze (ESSM) (12). After receiving mTESE tissue or a double ejaculate, multiple embryologists will painstakingly search these samples for over 6–8 hours. If found, individual sperm are isolated and used either fresh for intracytoplasmic sperm injection (ICSI) or preserved for later use. Reported recovery rates range from 40% to 44%.

A significant limitation to identifying rare sperm for use in ICSI is that selection must be nondestructive. Sperm specific biomarkers (e.g., labeled antibodies) cannot be used. Nor can traditional DNA fragmentation, chromatin integrity, or aneuploidy assays as these will requisitely damage sperm DNA. Advances in nondestructive imaging modalities such as RAMAN spectroscopy and holographic imaging have been investigated (13, 14). However, these platforms generally provide population-level information such as DNA fragmentation, protein oxidation, lipid damage, motility, and morphology, and are not geared toward rare sperm identification.

Enter artificial intelligence (AI). Ghadya et al. (15) provide a comprehensive review of the use of AI in male infertility; however, AI's use for the study of NOA remains limited. Emerging work has demonstrated its ability to detect rare sperm in dissociated mTESE samples using deep learning (16), machine learning to predict sperm extraction outcomes in patients with NOA (17), and use of Convolutional Neural Network (CNN) driven image analysis to identify rare sperm in mTESE samples (18). In the lattermost study, CNN-guided sperm search vastly outperformed trained embryologists in time to sperm identification (1,800× faster) with improved recall while maintaining a similar number of sperm identified (SpermSearchAI). Most recently, reproductive endocrinologists have reported the first successful pregnancy using sperm identified by AI-guided image analysis which were then sorted by microfluidics (sperm tracking and recovery (STAR) system) (19).

Herien, we introduce a novel diagnostic platform that addresses the challenge of identifying ultrarare sperm in ejaculate samples. Specifically, we paired high-throughput imaging cytometry with AI-assisted CNN guided image analysis to identify ultrarare sperm in ejaculate samples deemed clinically azoospermic.

MATERIALS AND METHODS

Sample processing, acquisition, and IRB approval

Our CLIA-approved andrology laboratory performs semen analyses following the WHO 6th edition guidelines (20) and processes surgically retrieved testicular tissue. If a neat ejaculatory sample is found without sperm on initial examination, the sample is centrifuged at 700 g for 10 minutes, the supernatant is aspirated, and the sample is resuspended in 50 μL of human tubal fluid media (InvitroCare, Frederick, MD). Four 10 μL aliquots are then examined individually, and the number of motile and immotile sperm is recorded. A sample is deemed azoospermic if no sperm are identified or cryptozoospermic if sperm are identified.

Surgically retrieved testicular tissue from mTESE samples is processed as follows. During the operative mTESE procedure, and before sample transfer to the corresponding in vitro fertilization (IVF) laboratory for ICSI, small representative sections of each numbered specimen are gently teased apart using two glass slides and examined under a phase-contrast microscope by the primary surgeon. No enzymatic digestion or tissue staining is performed. The presence or absence of mature sperm in each numbered sample is noted, and this information is communicated to the collaborating embryologist.

After these clinical analyses are performed, there often exists a remnant ejaculatory or testicular tissue specimen. If a sample is deemed of clinical interest, the sample is deidentified and transferred to our research biobank (IRB H-47408). Biobanked samples include either remnant 10 μL aliquots from pelleted ejaculatory samples or remnant testicular tissue samples analyzed during the operative mTESE (which are separate from the mTESE samples delivered to the collaborating IVF laboratory). This research protocol was deemed to be of minimal risk and did not require informed consent because all remnant biobanked samples were deidentified.

Samples were stored at 4 °C or cryopreserved at −196 °C. Samples deemed clinically azoospermic or cryptozoospermic were identified from the biobank. Controls included severely oligospermic (<5 M/mL), oligospermic (>5 M/mL but <15 M/mL), and normospermic (>15 M/mL) samples. Deidentified clinical data were obtained through chart review. This study was approved by our institution (IRB H-57730).

Sample preparation

Remnant ejaculatory samples were fixed and permeabilized (eBioscience Foxp3 / Transcription Factor Buffer Set, ThermoFisher Scientific, Waltham, MA) (21). Samples were stored at 4 °C until batched analyses. Samples were first stained with DAPI (ThermoFisher Scientific), rat anti-human CD45-PE (monoclonal YAML504.1, ThermoFisher Scientific), mouse anti-human acrosomal vesicle protein 1 (ACRV-1) (monoclonal 01, SinoBiological, Houston, TX), and rabbit anti-human lactate dehydrogenase C (LDHC) (monoclonal 9F2, ThermoFisher Scientific) (8). Samples were then counterstained with highly cross-absorbed secondary antibodies: goat anti-rabbit Alexa 488 (ThermoFisher Scientific) and goat anti-mouse Alex647 (ThermoFisher Scientific). Control samples were stained without the primary antibodies or mixed with peripheral blood mononuclear cells (PBMCs) to confirm antibody specificity. Cryopreserved testicular tissue was thawed to 37 °C, minced gently, and then incubated in a phosphate-buffered saline (PBS)/Collagenase IV Solution (2 mg/mL)/DNaseI (20 U/mL) solution at 37 °C for 45 minutes with gentle vortexing (22). The reaction was

then quenched, the tissue filtered through a cell strainer, fixed/permeabilized, and stained with the above reagents.

Imaging cytometry

Validation of antibody staining was performed using the Cytek Amnis ImageStream^{X} Mark^{II} imaging cytometer (8). Brightfield images of sperm and PBMCs were acquired at ×60 optical magnification (0.3 μm/pixel). Brightfield/fluorescent overlays were created using the 405 nm, 488 nm, 561 nm, and 642 nm laser lines. For high-throughput cytometric image acquisition, we utilized the ThermoFisher Attune Cytpix acoustic-focusing cytometer, capable of imaging 6,000 brightfield events per second at a similar resolution of 0.3 μm/pixel (248 × 248 pixels). Although this instrument does not provide brightfield/fluorescent image overlays, each imaged event contains corollary fluorescent data. Samples were grouped by acquisition date to track potential run-to-run variation. Six batched runs were conducted. Images were exported as 8-bit TIFF files and fluorescence data as FCS 3.1 files. Images and FCS files were analyzed using our self-programmed graphical user interface (GUI). All data were stored securely on a remote BCM server.

Model architecture

A ConvNeXt CNN model pretrained on ImageNet-1k was used as the base model for image classification (23). The ConvNeXt architecture was designed to emulate the higher accuracy achievable with large transformer-based models while maintaining the efficiency of standard CNNs and is well-suited to the classification of large-scale biomedical image datasets (24). The classification head was replaced with a two-unit fully connected layer for binary classification. The entire network was fine-tuned using cross-entropy loss and the AdamW optimizer with a cosine-annealing learning rate schedule and periodic warm restarts on a manually labeled training dataset selected from the first five batched runs using an in-house Python-based GUI.

Cross validation and test set validation

Model training and evaluation on the training dataset followed five-fold stratified cross-validation. For each fold, 20% of sperm and debris images were reserved as a validation set, and the remaining 80% formed the training set. To ensure balanced classes during training, debris images were randomly subsampled until equal numbers of sperm and debris were included in the training data. The model was then trained, and performance was assessed on the corresponding validation set. The process was repeated for five unique folds to generate nonoverlapping test partitions. A final model was then trained on all training set images and assessed using a separate test set comprising images manually selected and labeled from the sixth batch run.

Comprehensive sperm search and review

All images were classified as sperm or debris by the model. These classifications were saved to a CSV file along with the batch run, sample number, and unique image identifier for tracking and later analysis. To provide an unbiased assessment of model performance in practice, all images from azoospermic patients classified as sperm by the model were manually verified as sperm or debris by three independent reviewers, and conflicting annotations were assigned the label corresponding to the majority vote. Positive events imaged were confirmed to express the sperm-specific markers ACRV1 and LDHC using our self-programmed GUI (Supplemental Fig. 1, available online). Positive predictions in samples from cryptozoospermic and severely oligospermic men were validated by a single reviewer until 100 sperm had been verified or all images in the sample had been reviewed.

Statistical analyses

Performance of the model during cross-validation and test set validation was assessed using F1-score, sensitivity, and specificity. Interrater agreement for the manual review of positive model predictions in the azoospermic samples was assessed using Cohen's Kappa. Equations for the performance measures are given in Supplemental Figure 2. All model training, validation, and testing were performed using Python version 3.13 with the PyTorch, sklearn, and timm libraries (25). Statistical analyses were conducted using the sklearn and numpy libraries.

RESULTS

Samples identified

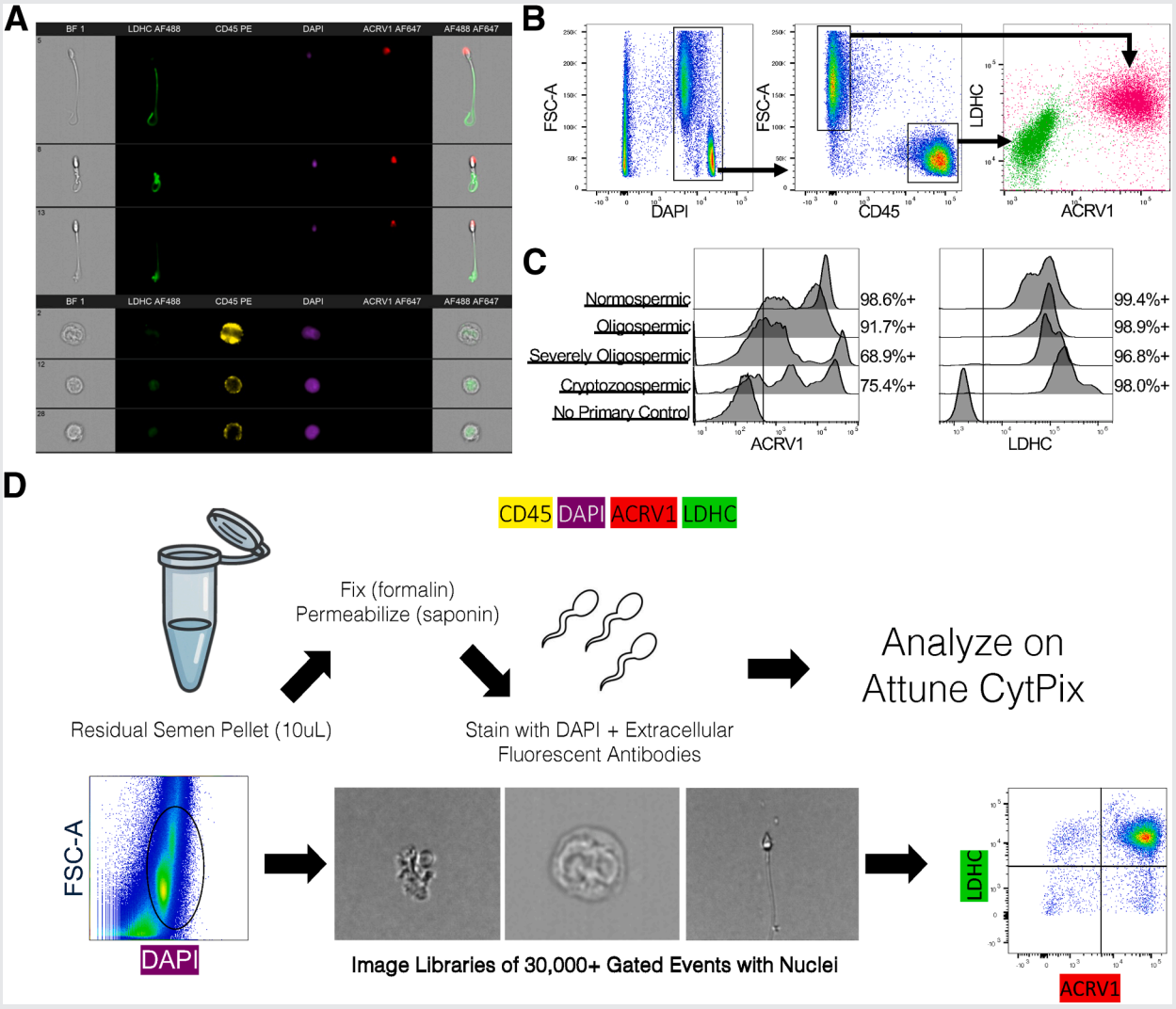
From our biobank, we identified 83 azoospermic, 34 cryptozoospermic, and 22 severely oligospermic remnant samples. For model training and validation, we also identified five oligospermic and four normospermic remnant samples. Samples were identified from December 2024 to September 2025. Presumed etiologies of azoospermia were extracted from electronic charts.

Verification of staining protocol

To verify our sperm-specific staining protocol for use in imaging cytometry, we built upon work by Zhang et al (8) utilizing sperm-specific antibodies targeted to ACRV1 and LDHC. We mixed a normospermic semen sample with PBMCs, which served as a negative antigenic control. This mixed sample was then stained with DAPI (a DNA marker), ACRV1, LDHC, and CD45 (protein tyrosine phosphatase C, which is expressed by cells of the hematopoietic lineage) and imaged using a Cytek Amnis ImageStream^{X} Mark^{II} (ImageStream) imaging cytometer. The DAPI was positive in all cells, ACRV1 was specific to sperm heads, and CD45 was specific to mononuclear cells (Fig. 1A). The LDHC localized primarily to sperm tails with high expression.

Titration of these antibodies found ACRV1 and LDHC expression to be specific to sperm when mixed with PBMCs (Fig. 1B, pink cells CD45 negative, green cells CD45 positive). Images were acquired at the highest resolution (×60 magnification, 1 × 1 binning) with a median rate of 5.7 ± 0.3 images/s. We then stained representative cryptozoospermic, severely

FIGURE 1



Development of a sperm-specific antibody staining protocol for high-throughput imaging cytometry. (**A**) Conventional imaging cytometry demonstrates ACRV1 expression on sperm heads and LDHC expression on sperm tails, which were absent on CD45-positive leukocytes. (**B**) Fluorescent plots of an experiment in which a normospermic ejaculate sample was mixed with PBMCs. After selecting for nuclei-positive events (DAPI+, *left* panel), one can separate CD45+ PBMCs from CD45- sperm (*middle* panel), and then assess expression of ACRV1 and LDHC (*right* panel). Sperm (*pink*) were ACRV1 and LDHC double positive, and PBMCs (*green*) were ACRV1 and LDHC double negative. (**C**) Expression of ACRV1 and LDHC based on ejaculate sample concentration. LDHC expression appeared consistent regardless of sample concentration, whereas ACRV1 expression varied. (**D**) Sample preparation protocol for analysis on the high-throughput imaging cytometer (*top* panel). The output data sets of DAPI+ events analyzed included individual raw 8-bit TIFs with matched fluorescent data (*bottom* panel, normospermic data set shown). ACRV1 = acrosomal vacuole protein 1; LDHC = lactate dehydrogenase C; PBMCs = peripheral blood mononuclear cells.

Stocks. Imaging cytometry redefines azoospermia. Fertil Steril 2026.

oligospermic, oligospermic, and normospermic samples vs. an unstained, no primary antibody control (Fig. 1C) to further confirm marker specificity across sample type. The ACRV1 expression level varied on the basis of the concentration of the ejaculate, consistent with previous work (26), but LDHC expression appeared consistent. With confidence in this staining protocol, we transitioned to data collection using the high-throughput imaging cytometry platform.

High throughput imaging cytometry

The Attune Cytpix imaging cytometer uses acoustic waves to center particles in a single file line, enabling precise acquisition of fluorescent and imaging data even at high flow rates. Using the staining protocol described (Fig. 1D), we analyzed remnant ejaculatory samples from azoospermic (n = 83), cryptozoospermic (n = 34), and severely oligospermic (n = 22) men in six batched runs (Supplemental Table 1,

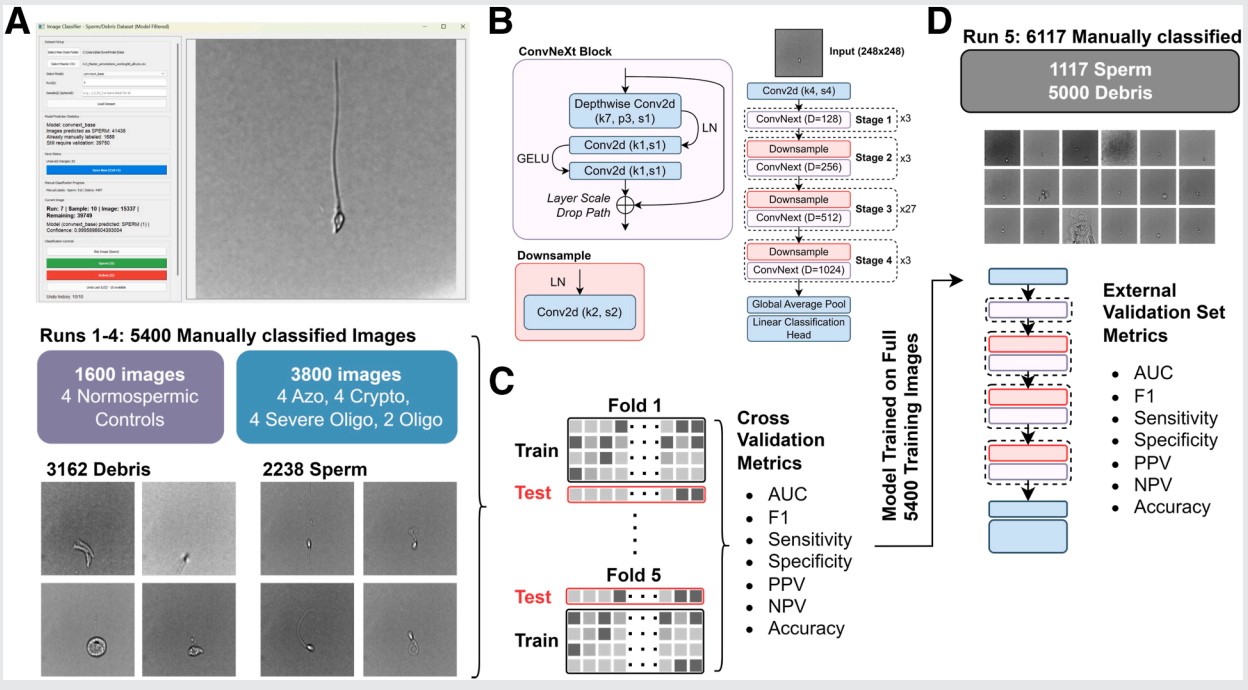
available online). We also imaged oligospermic (n = 5) and normospermic (n = 4) controls, which would later serve for model training and cross-validation. Total images captured were 2,164,907 (azoospermic), 1,023,289 (cryptozoospermic), and 62,0971 (severely oligospermic). Including oligospermic and normospermic samples, 4,085,172 unique images across 148 samples were collected (Supplemental Table 1). Median images captured per sample were 30,000 (interquartile range [IQR] 22,883–30,000), which is the upper limit per sample set by the imaging cytometer. Median rate of image acquisition was 177.5 images/s (IQR 77.9–285.7), representing a 31-fold increase over conventional imaging cytometry. Median individual sample acquisition time was 150 seconds (IQR 104–283) and overall median fluorescent cellular event rate per second was 7,921 (IQR 3,129–38,320).

Before proceeding with data collection, we tested rates of sample spillover using different washing techniques by running a normospermic sample followed by an "empty" water sample. This was important in that we detected a 48% cell event spillover rate when no cleaning protocol was performed. After testing multiple protocols, we found that a complete cytometer sanitize with 50% Contrad cleaner, followed by a complete cytometer sanitize with water, yielded a potential cell spillover rate of just 0.02% between a normospermic sample and the subsequent water sample. Therefore, by employing this safety protocol, any potential spillover from an azoospermic sample to a subsequent azoospermic sample would be obviated.

CNN model training, validation, and predictions

The initial training dataset was comprised of 5,400 images drawn from the first five runs, encompassing normospermic controls and from patients across the azoospermia-oligospermia spectrum (Fig. 2A). Each of the 5,400 images was manually classified as sperm (n = 1,600) or debris (n = 3,800) using an in-house GUI. Model architecture is depicted in Figure 2B and cross-validation is depicted in Figure 2C. To evaluate generalizability across potentially variable acquisition conditions, an additional annotated set consisting of 6,117 images (1,117 sperm and 5,000 debris) from the sixth run was prepared for validation (Fig. 2D).

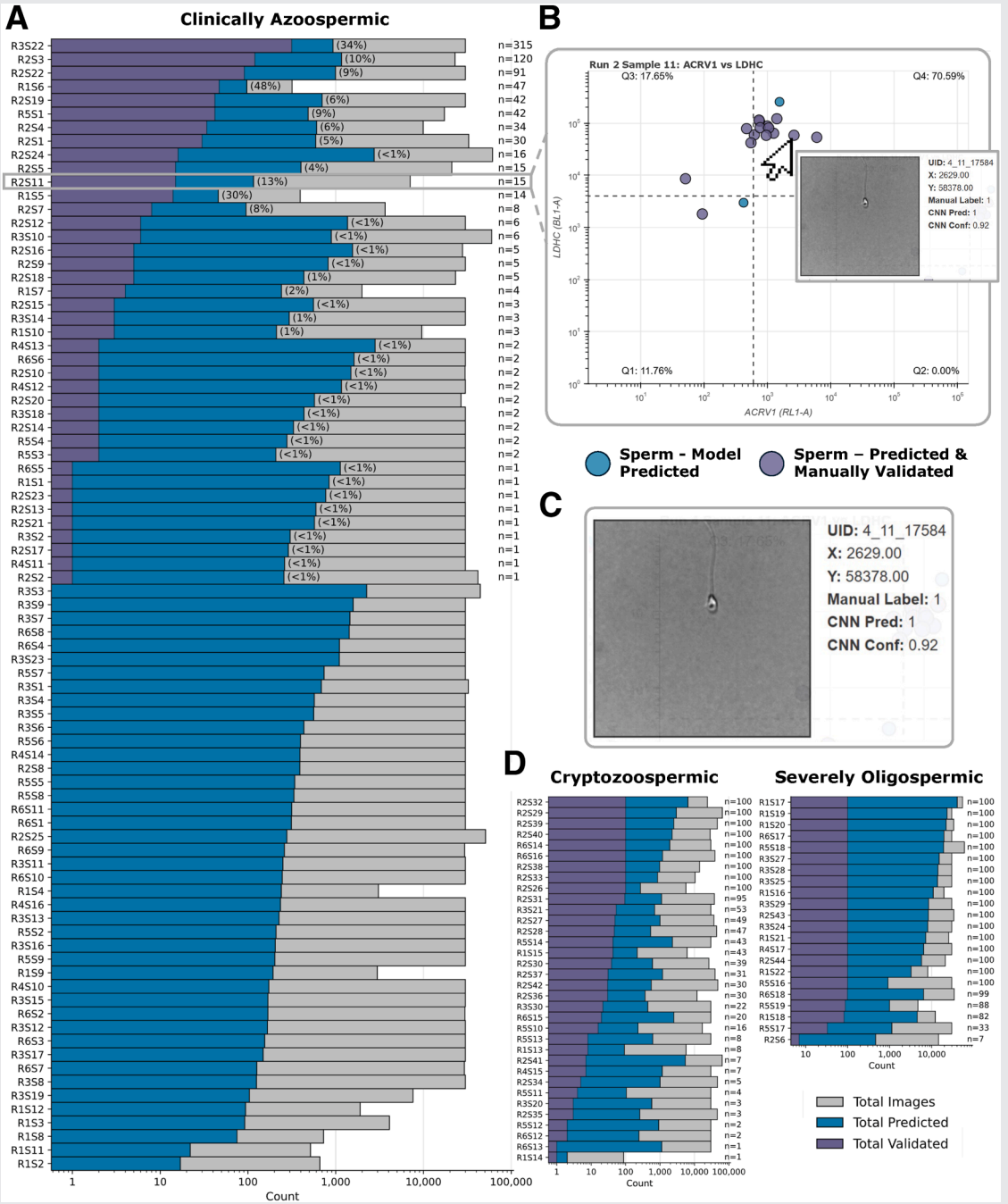
FIGURE 2



Model architecture, training, and validation. (**A**) Manual classification user interface. As images are visually classified as sperm or debris, an annotation file for model training is automatically generated. The training set comprises images from the first five runs. (**B**) ConvNeXt model architecture. The stem divides the images into patches of 4 × 4, which are passed through four stages of progressively increasing channel dimension and decreasing spatial resolution. The output is averaged and passed through a final fully connected head for binary classification. (**C**) Fivefold cross-validation. For each fold, 20% of the sperm and debris images are set aside as a validation set whereas the remaining 80% form the training set. Debris images are randomly excluded from the training set until there is an equal number of sperm and debris in the training set to ensure balanced training. The model is then trained, and the performance is measured on the validation set. The process is repeated four more times, drawing from a unique 20% of the sperm and debris sample pools each time to create nonoverlapping validation sets. (**D**) A model trained on all 5,400 training set images is used to classify 6,117 images from Run 6 manually annotated as sperm or debris. This test-set validation evaluated the performance of the model on images acquired from a run not used in initial training. AUC = area under the receiver operating characteristic curve; Azo = azoospermic; Crypto = cryptozoospermic; GELU = Gaussian error linear units; k = kernel size; LN = LayerNorm; NPV = negative predictive value; Oligo = oligospermic; p = padding; PPV = positive predictive value; s = stride.

Stocks. Imaging cytometry redefines azoospermia. Fertil Steril 2026.

FIGURE 3



Predictions of the automated sperm detection model. (**A**) Log-scale bar plot of total images captured for each sample in men with prior diagnosis of azoospermia (*gray*), total images classified as sperm by the model (*blue*), and total images visually verified as sperm by the three independent reviewers (*purple*). Percent of model predictions manually verified as sperm is shown to the right of the corresponding *blue bar*, whereas total sperm verified for each sample is shown to the right of the *gray bar*. (**B**) Expression levels of ARCV1 acrosomal protein plotted against LDHC sperm-specific lactate dehydrogenase isozyme in putative sperm cells for a representative sample. (**C**) Demonstration of graphical user interface. The user can view images captured by the imaging flow cytometer by hovering over the location of corresponding marker expression and search for specific images by ID-number (not shown). (**D**) Log-scale bar plots of total images (*gray*), total images classified as sperm by the model (*blue*), and total images visually verified as sperm by an independent review (*purple*) in cryptozoospermic and severely oligospermic samples. **Bolded brackets** indicate run, whereas plain-font labels indicate sample number within the run. Manual validation was stopped once 100 sperm were identified in the cryptozoospermic and severely oligospermic samples.

Stocks. Imaging cytometry redefines azoospermia. Fertil Steril 2026.

The ConvNeXt model achieved an F1-score of 0.963 ± 0.006, sensitivity of 0.971 ± 0.007, and specificity of 0.955 ± 0.010 during the fivefold cross-validation on the initial 5,400-image training dataset. When evaluated using the 6,117-image test set, the model achieved an F1-socre of 0.874, sensitivity of 0.881, and specificity of 0.970 (Supplemental Fig. 3). The model classified at least one image as sperm in all azoospermic samples (Fig. 3A). Manual validation found sperm in 40/83 (48%) of these samples (see Supplemental Fig. 1 for a full graphical depiction of our manual validation process). All three reviewers agreed on a label in 94.3% of the verified images, resulting in pairwise Cohen's Kappa ranging from 0.723 to 0.907 between reviewers. Sperm were identified by the model in all cryptozoospermic and severely oligospermic samples; manual validation confirmed at least one sperm in all 34 cryptozoospermic and all 22 severely oligospermic samples.

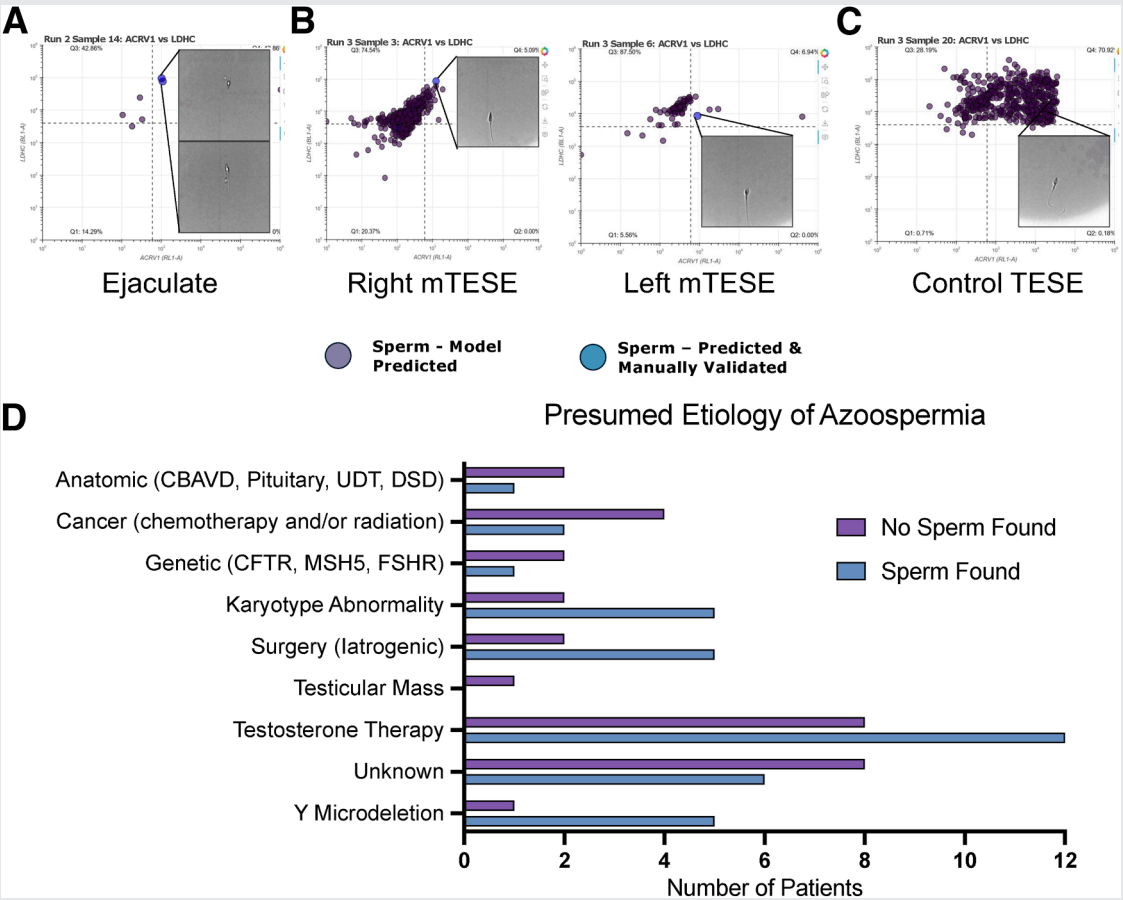
CNN training and inference speed

Fine-tuning all weights of the modified ConvNeXt classifier on the 5,400-image training set required 20.1 minutes of GPU time, 1.62 GB of vRAM, and 1.72 GB of RAM on average across the five folds. During inference, the model classified images at an average rate of 265 images/s, requiring 490 MB vRAM and 2.87 GB RAM. A full sample with the median of 30,000 images requires 113 seconds of total processing time to classify.

Corollary mTESE surgical specimen

One clinically azoospermic individual had two unique remnant ejaculatory samples from which our CNN model predicted sperm which we confirmed manually via our GUI (Fig. 4A, example images with corollary ACRV1/LDHC positivity on the fluorescent plot). During the period assessed,

FIGURE 4



CNN model predicts sperm in a negative mTESE sample and clinical phenotypes. (**A**) Representative GUI with manually validated sperm found in an azoospermic ejaculate sample. (**B**) This same individual underwent a subsequent bilateral mTESE in which no sperm were identified by the surgeon or embryologist. The CNN model predictions of right and left mTESE samples with manually validated sperm. (**C**) Sperm were identified from a control TESE sample from an individual status after vasectomy. (**D**) Presumed primary etiologies of azoospermia in samples analyzed. Etiologies of azoospermia were classified as y-microdeletion, unknown, testosterone therapy, testicular mass, surgery (iatrogenic), karyotype abnormality, genetic, cancer with exposure to chemotherapy/radiation, or anatomic. Samples with CNN-predicted and manually validated sperm were then compared (sperm found – *blue bars*; no sperm found – *purple bars*).

Stocks. Imaging cytometry redefines azoospermia. Fertil Steril 2026.

this individual underwent a bilateral mTESE in which no sperm were seen by the surgeon or embryologist. From our biobank, these right and left remnant mTESE tissue specimens were analyzed in triplicate. Our CNN predicted sperm in both left (1 of 3) and right (2 of 3) replicate tissue samples, which were confirmed manually by visual morphologic confirmation and fluorescent expression (Fig. 4B). As a control, our CNN model also reliably found sperm from a TESE sample from an individual post vasectomy (Fig. 4C).

Clinical etiologies of azoospermia

We extracted a list of presumed etiologies of azoospermia from the 83 ejaculate samples analyzed. Of the 67 unique specimens deemed clinically azoospermic, 37 had manually validated sperm from the CNN model predictions and 30 did not. We grouped presumed etiologies of azoospermia into nine categories and compared the "sperm found" and "no sperm found" groups (Fig. 4D). The presumed etiologies of azoospermia did not appear to affect whether our CNN predicted the presence or absence of sperm.

DISCUSSION

We introduce a novel diagnostic platform that pairs high-throughput imaging cytometry with AI-assisted image analysis to identify ultrarare sperm in azoospermic samples. We developed a fluorescent-antibody staining protocol to confirm expression of sperm-specific proteins on imaged events that appeared morphologically congruent with sperm. We programmed and cross-validated a CNN model to identify sperm with high sensitivity and specificity, which were manually validated using an internally developed GUI. Overall, we report a 48% detection rate of ultrarare sperm in remnant azoospermic ejaculate samples and identified sperm from an mTESE sample in which none were found by the operating surgeon or embryologists.

Key limitations of our dataset warrant consideration. One, the samples analyzed were remnant specimens capturing approximately one-fifth of the clinical analyte. We purposefully utilized our retrospective biobank to optimize analytical throughput. Efforts are currently underway using a separately approved institutional review board (IRB) protocol (H-57193) that permits the collection of whole semen samples for prospective research use after informed consent. Two, our experimental staining protocol could not be used for clinical ICSI. Staining for DNA (DAPI) and the intracellular sperm-specific proteins ACRV-1 and LDHC requires cellular fixation and permeabilization. This staining protocol was developed to confirm whether an imaged event morphologically congruent with a mature sperm also expressed sperm-specific proteins. Three, samples were not run in technical nor biologic replicate because of the study design and execution. Four, as with any rare cell analysis, avoiding cellular spillover between samples will be essential, especially for future clinical diagnostic use. We took great care to mitigate this risk in experimental design and acquisition. Five, the upper limit of image acquisition per sample was 30,000 events. Exhaustive runs may be needed for complete analysis of prospectively collected samples. Six, this platform lacks data on sperm motility as only static images were analyzed. Seven, as the specific imaging cytometer utilized does not sort nor recover individual cells, we could not assess the reproductive capacity of the identified sperm.

The extremely low sperm concentration in our target population poses a substantial challenge for training the CNN classifier. High specificity is paramount, because even the 95% specificity reached by our model corresponds to approximately 1,500 false positives for a 30,000-image sample. Nonetheless, sensitivity must also remain high to detect the rare sperm in these men. Given the few sperm visible in azoospermic men, it is also difficult to create a robust training set for the model, because there are simply not enough positive events in our population of interest. Consequently, we included normospermic, oligospermic, and cryptozoospermic men to supplement training, although this may bias the model because of subtle differences in debris composition and sperm morphology. Image resolution is another limitation; at only 248 $\times$ 248 pixels (0.3 μm/pixel), distinguishing sperm from debris is often difficult even for human experts, as reflected by the interreviewer Cohen's Kappa. Larger models with deeper architectures can improve classification performance to a certain degree, but this comes at the cost of greater inference time required to call an image as sperm or debris.

Finally, significant validation will be required before the clinical introduction of our novel diagnostic platform. Although other state-of-the-art platforms for rare sperm search are emerging (ESSM, SpermSearchAI, and STAR system), their widespread clinical adoption also remains limited. Direct head-to-head comparisons of these platforms will be necessary, and significant coordination and collaboration will be required between research groups.

Looking to the future, high-throughput imaging cytometry paired with AI-driven image analytics could be extended from diagnostic to therapeutic use. Newer generations of imaging flow cytometers now combine real-time, high-speed cell imaging with gentle cell-sorting. One could envision running an untouched ejaculate or digested mTESE sample through these machines, and simply, through image acquisition and reconstruction (without using any destructive staining protocols), be sorted for sperm based on morphology alone. Sorted rare sperm could then be used for ICSI or cryopreserved. Although significant preclinical testing and validation will be required before clinical use (including testing of reproductive capacity), this platform may represent a means to identify and collect ultrarare sperm in ejaculate or mTESE samples.

We propose a new sperm concentration, microzoospermia, a classification falling between true azoospermia and classically defined cryptozoospermia. We define microzoospermia as "a sperm concentration so low that even after manual analysis of a pelleted sample finds no sperm, only advanced imaging-based search technologies can reliably identify ultrarare sperm". Ideally, patients deemed microzoospermic may one day avoid mTESEs altogether using these emerging technologies. Future studies will validate the data presented herein through prospective sample collection, biological and technical replication, as well as CNN model refinement.

CONCLUSION

High-throughput imaging cytometry paired with AI-assisted analysis enabled the identification of ultrarare sperm in clinically azoospermic ejaculate samples. We report a 48% detection rate and propose "microzoospermia" as a novel sperm concentration category identifiable only through advanced imaging and search technologies.

CRediT Authorship Contribution Statement

Blair Stocks: Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization. **Aidan Boyne:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation. **Amelia Oppenheimer:** Writing – review & editing, Data curation. **Katelynn Salmon:** Writing – review & editing, Formal analysis. **Taylor Kohn:** Writing – review & editing, Conceptualization. **Gal Saffati:** Writing – review & editing, Data curation. **Beatriz Hernandez:** Writing – review & editing, Data curation. **Justin Badal:** Writing – review & editing, Formal analysis. **Mohit Khera:** Writing – review & editing, Conceptualization. **Larry Lipshultz:** Writing – review & editing, Conceptualization.

Declaration of Interests

B.S. has nothing to disclose. A.B. has nothing to disclose. A.O. has nothing to disclose. K.S. has nothing to disclose. T.K. has nothing to disclose. G.S. has nothing to disclose. B.H. has nothing to disclose. J.B. has nothing to disclose. M.K. serves as a consultant for Endo Pharmaceuticals, Boston Scientific, Verity Pharmaceuticals, Marius Pharmaceuticals, and Besins Pharmaceuticals. L.L. serves as a consultant for AbbVie, Arex Life Sciences, Contraline, Endo Pharmaceuticals, Lipocine, and Teladoc. L.L. serves as an advisor for Inherent Biosciences. L.L. is a speaker for Boston Scientific.

SUPPLEMENTAL MATERIAL

Supplemental data for this article can be found online at https://doi.org/10.1016/j.fertnstert.2026.04.011.

REFERENCES

1. Brannigan RE, Hermanson L, Kaczmarek J, Kim SK, Kirkby E, Tanrikut C. Updates to male infertility: AUA/ASRM guideline (2024). J Urol 2024;212:789–99.
2. Bernie AM, Mata DA, Ramasamy R, Schlegel PN. Comparison of microdissection testicular sperm extraction, conventional testicular sperm extraction, and testicular sperm aspiration for nonobstructive azoospermia: a systematic review and meta-analysis. Fertil Steril 2015;104:1099–103.e1–3.
3. Corona G, Minhas S, Giwercman A, Bettocchi C, Dinkelman-Smit M, Dohle G, et al. Sperm recovery and ICSI outcomes in men with non-obstructive azoospermia: a systematic review and meta-analysis. Hum Reprod Update 2019;25:733–57.
4. Achermann APP, Pereira TA, Esteves SC. Microdissection testicular sperm extraction (micro-TESE) in men with infertility due to nonobstructive azoospermia: summary of current literature. Int Urol Nephrol 2021;53:2193–210.
5. Han TX, Berk B, Ghayda RA, Ernandez J, Kathrins M. Financial decision analysis based on "willingness to pay" for surgical sperm retrieval approaches among men with non-obstructive azoospermia in the United States. Andrology 2024;12:422–8.
6. Stocks B, Lipshultz L, Brogaard K, Kohn T, Oppenheimer A, Peterson D, et al. PD23-04 A non-invasive predictor of sperm retrieval success in men with NOA: Single-molecular native methylation sequencing for sperm-specific cfDNA. J Urol 2025;213:e825.
7. Anbazhagan R, Rivas R, Tan K, Hsieh M, Capponi C, Bastuba M, et al. A noninvasive diagnostic assay that detects spermatogenic defects and single sperm. Andrology 2025;13:118–9.
8. Zhang J, Kanoatov M, Jarvi K, Gauthier-Fisher A, Moskovtsev SI, Librach C, et al. Germ cell-specific proteins AKAP4 and ASPX facilitate identification of rare spermatozoa in non-obstructive azoospermia. Mol Cell Proteomics 2023;22:100556.
9. Ghanami Gashti N, Sadighi Gilani MA, Jabari A, Qasemi M, Feizollahi N, Abbasi M. The germ cell-specific markers ZPBP2 and PGK2 in testicular biopsies can predict the presence as well as the quality of sperm in non-obstructive azoospermia patients. Reprod Sci 2021;28:1466–75.
10. Jarvi K, Schlegel P, Schiza C, Drabovich A, Lau S, Soosaipillai A, et al. Semen biomarker TEX101 predicts sperm retrieval success for men with testicular failure. F1000Research 2021;10:569.
11. Ceyhan E, Kayra MV, Gul Ates E, Kizilkan Y, Altan M, Yildirim O, et al. A Nomogram predicting testicular sperm extraction success in men with non-obstructive azoospermia: a multi-center study. Urology 2025;196:155–61.
12. Miller N, Biron-Shental T, Pasternak Y, Belenky M, Shefi S, Itsykson P, et al. Fertility outcomes after extended searches for ejaculated spermatozoa in men with virtual azoospermia. Fertil Steril 2017;107:1305–11.
13. Zhang F, Tan Y, Ding J, Cao D, Gong Y, Zhang Y, et al. Application and progress of Raman spectroscopy in male reproductive system. Front Cell Dev Biol 2022;9:823546.
14. Wheeler MB, Rabel RAC, Rubessa M, Popescu G. Label-free, high-throughput holographic imaging to evaluate mammalian gametes and embryos†. Biol Reprod 2024;110:1125–34.
15. Ghayda RA, Cannarella R, Calogero AE, Shah R, Rambhatla A, Zohdy W, et al. Artificial intelligence in andrology: from semen analysis to image diagnostics. World J Mens Health 2024;42:39–61.
16. Lee R, Witherspoon L, Robinson M, Lee JH, Duffy SP, Flannigan R, et al. Automated rare sperm identification from low-magnification microscopy images of dissociated microsurgical testicular sperm extraction samples using deep learning. Fertil Steril 2022;118:90–9.
17. Bachelot G, Dhombres F, Sermondade N, Haj Hamid R, Berthaut I, Frydman V, et al. A machine learning approach for the prediction of testicular sperm extraction in nonobstructive azoospermia: algorithm development and validation study. J Med Internet Res 2023;25:e44047.
18. Goss DM, Vasilescu SA, Vasilescu PA, Cooke S, Kim SH, Sacks GP, et al. Evaluation of an artificial intelligence-facilitated sperm detection tool in azoospermic samples for use in ICSI. Reprod Biomed Online 2024;49:103910.
19. Suryawanshi H, Gemmell LC, Morgan S, Koustas G, Prosser RW, Fu R, et al. First clinical pregnancy following AI-based microfluidic sperm detection and recovery in non-obstructive azoospermia. Lancet 2025;406:2213–4.
20. World Health Organization. WHO laboratory manual for the examination and processing of human semen. 6th ed. Geneva: World Health Organization; 2021.
21. Stocks B, Oppenheimer A, Tran L, Huang J, Lipshultz L. 186 The use of the fluorescent nuclear stain DAPI to identify seminal leukocytes. J Sex Med 2024;21, qdae167.181.
22. Barrachina F, Ottino K, Elizagaray ML, Gervasi MG, Tu LJ, Markoulaki S, et al. Regulatory T cells play a crucial role in maintaining sperm tolerance and male fertility. Proc Natl Acad Sci USA 2023;120:e2306797120.
23. Liu Z, Mao H, Wu C-Y, Feichtenhofer C, Darrell T, Xie S. A ConvNet for the 2020s. arXiv 2022;arXiv:2201.03545.
24. Mmileng OP, Whata A, Olusanya M, Mhlongo S. Application of ConvNeXt with transfer learning and data augmentation for malaria parasite detection in resource-limited settings using microscopic images. Plos One 2025;20:e0313734.
25. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: an imperative style, high-performance deep learning library. arXiv 2019;arXiv:912.01703.
26. Coppola MA, Klotz KL, Kim K-A, Cho HY, Kang J, Shetty J, et al. SpermCheck Fertility, an immunodiagnostic home test that detects normozoospermia and severe oligozoospermia. Hum Reprod 2010;25:853–61.

La citometría de imagen de alto rendimiento combinada con inteligencia artificial identifica espermatozoides ultrarraros en hombres previamente diagnosticados con azoospermia clínica

Objetivo: Determinar si una nueva plataforma diagnóstica, que combina citometría de imagen de alto rendimiento con análisis de imágenes asistido por inteligencia artificial (IA), puede identificar espermatozoides ultrarraros en muestras eyaculatorias consideradas clínicamente azoospérmicas.

Diseño: Revisión retrospectiva de cohorte de muestras eyaculatorias remanentes biobancadas procesadas mediante citometría de imagen de alto rendimiento y análisis de imágenes asistido por IA.

Sujetos: Muestras eyaculatorias remanentes (n = 148) biobancadas entre diciembre de 2024 y septiembre de 2025 en una clínica de infertilidad masculina de alto volumen.

Exposición: Citometría de flujo basada en imágenes y análisis de imágenes impulsado por redes neuronales convolucionales (CNN) para la clasificación diagnóstica.

Principales medidas de resultado: Tasa de reclasificación diagnóstica de muestras eyaculatorias remanentes inicialmente consideradas clínicamente azoospérmicas mediante evaluación morfológica basada en CNN y expresión correlativa de proteínas específicas de espermatozoides.

Resultados: Entre las 83 muestras azoospérmicas analizadas, se encontraron espermatozoides en 40 muestras (48%) utilizando esta plataforma diagnóstica de alta fidelidad.

Conclusión: Proponemos una nueva clasificación de concentración espermática, microzoospermia, definida como una concentración de espermatozoides tan baja que solo las tecnologías avanzadas de búsqueda basadas en imágenes pueden identificar de manera confiable espermatozoides ultrarraros.