

Development and clinical evaluation of a microfluidic device to automate and improve oocyte recovery from follicular fluid in patients undergoing in vitro fertilization

Corresponding Author: Dr Emre Ozkumur

A version of this paper was originally rejected for publication by Nature Medicine, however that decision was reconsidered after appeal by the authors.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I read this study with great interest. The authors, for the first time, report a microfluidic device (FIND-Chip) that automates the isolation and denudation of oocytes from follicular fluid (FF) for IVF. In addition to validation using a bovine model, they take an important step forward by demonstrating the device's potential clinical utility, successfully identifying otherwise discarded oocytes in approximately half of the samples tested from ~500 patients. They further support this with a small pilot study showing that these recovered oocytes can contribute additional blastocysts to the embryo cohort.

The manuscript is well written, and the experiments follow a clear and logical trajectory. I believe this technology represents a meaningful advance in IVF, with implications for improving success rates, workflow efficiency, and accessibility. I have only minor comments:

- Could the authors discuss next steps for evaluating this technology beyond the current pilot? For instance, would a randomized controlled trial (RCT) comparing standard embryologist-led screening versus the FIND-Chip, perhaps using the FF from the same patient, be feasible or informative?
- In IVF cycles triggered with a GnRH agonist, embryologists often visually confirm the presence of oocytes in the FF as an indirect check that the trigger worked. Do the authors envision incorporating an alert mechanism in the device to flag when an oocyte is identified? Alternatively, could the embryologist still briefly inspect the fluid before loading it into the device after the first egg is seen? A brief discussion on this could enhance the manuscript.
- Line 40: Recommend replacing "prognosis positive" with "good prognosis."; Line 345: The symbol for "less than or equal to 10" does not render properly, please correct the formatting; Lines 455–457: Please indicate whether a normality test was conducted prior to applying ANOVA.

Reviewer #2

(Remarks to the Author)

Comments to the Authors

Thank you for the opportunity to review this manuscript, which presents the microfluidic FIND-Chip designed to recover previously undetected oocytes from follicular fluid in IVF.

A. Summary of the Key Results

This manuscript reports findings from four distinct investigations using the FIND-Chip device:

1. Technology optimization using clinical samples spiked with bovine oocytes (N = 10 experiments with 20 oocytes each), demonstrating high isolation and denudation efficiency and equivalent or slightly improved embryo development rates.
2. Technology validation using donated human oocytes (N = 21 patients; 62 oocytes), showing 100% recovery of spiked-in oocytes and identification of additional "stealth" oocytes not detected by embryologists.
3. Multi-clinic investigation of discarded, pre-screened follicular fluid from IVF centers (N = 537 patients), where FIND-Chip identified extra oocytes in 54.9% of cases, regardless of the patient's original oocyte yield.

4. Clinical utilization pilot study (N = 19 patients), where recovered oocytes were returned to the clinical workflow, resulting in one ongoing pregnancy.

These are compelling early findings that suggest the system's promise for clinical use.

B. Originality and Significance

The work is novel and impactful. Demonstrating that a substantial proportion of oocytes are missed during standard IVF workflows and can be recovered by microfluidic automation addresses a longstanding challenge in reproductive medicine. Increasing the usable oocyte pool—particularly for low responders—has potential to improve patient outcomes and reduce the number of cycles needed.

C. Data & Methodology: Validity, Quality, Presentation

The methodology is generally sound, with well-documented optimization in bovine and validation in human oocytes. The large, multi-clinic dataset strengthens observational findings. However:

- Clinical investigations involving human oocytes are small, especially the pilot utilization study (N = 19, with only 12 mature extra oocytes used for ICSI). Conclusions about embryo competence and pregnancy potential should be framed as preliminary.
- Define the “fragmented or deformed” oocyte category (13.4% of extras) and clarify whether these were viable. Indicate whether atypical morphology affected device performance (e.g., clogging or failure).
- Acknowledge that only 17.9% of FF samples were processed in full volume, potentially underestimating extra-oocyte prevalence. Explain how partial volumes were selected and whether sampling was representative.
- Discuss possible reasons for inter-clinic variability in extra-oocyte rates (Table 1), including procedural or patient population differences.

D. Appropriate Use of Statistics & Uncertainty Management

Statistical comparisons are appropriate but underpowered in the pilot trial. Variability in fertilization and blastocyst rates should be addressed, and p-values provided where available. Clarify whether observed differences (e.g., 63% vs. 49% blastocyst rate) were statistically significant or descriptive only. Ensure all error bars are defined and statistical assumptions stated. Future studies should consider more robust models (e.g., mixed-effects regression) to account for patient- and clinic-level variability.

E. Conclusions: Robustness & Reliability

This is a strong and promising study introducing a potentially transformative technology in IVF. The technological performance is convincing, but clinical impact evidence is early-stage. The authors should clearly delineate the four studies, expand clinical validation in larger prospective trials, and address the above methodological and reporting points to maximize impact and suitability for Nature Medicine.

F. Suggested Improvements

1. The current title does not reflect the multi-phase nature of the work. Consider rephrasing to highlight both the technical development and the early clinical validation, e.g.: “Development and Early Clinical Evaluation of a Microfluidic System for Oocyte Recovery from Discarded Follicular Fluid”
2. Consider publishing the four distinct studies as stand-alone manuscripts for clarity and depth.
3. Provide a CONSORT-style diagram summarizing patient inclusion and clinical embryo outcomes.
4. Discuss the timeline, design, and endpoints for a prospective cohort or randomized trial to evaluate FIND-Chip's clinical impact (e.g., cumulative live birth rate).
5. Address FF sample handling time, device throughput, and operator variability.
6. Ensure all error bars and statistical comparisons are clearly defined.
7. Include embryo grading methodology and whether assessments were blinded.
8. Clearly distinguish modeled projections of live birth rate improvement from measured outcomes. In discussion section (lines 359-366), the study's projections of improved live birth rates are based on models, not measured clinical outcomes. Actual improvements remain unproven and should be presented as potential benefits pending further clinical validation.
9. Correct “pre-implentation genetic testing” to “pre-implantation genetic testing (PGT).”

G. References

The references are adequate and appropriate.

H. Clarity, Context & Policy Compliance

The manuscript is well written and coherent. The abstract effectively summarizes key findings.

Reviewer #3

(Remarks to the Author)

SUMMARY

This study provides proof of concept for a microfluidic device (FIND-Chip) that automates isolation and denudation of cumulus-oocyte-complexes (COCs) from follicular fluid (FF) aspirates obtained for IVF. Under experimental conditions, the device is shown to efficiently isolate and denude bovine and human COCs. Furthermore, re-screening of pre-screened FF

shows that the device identifies additional oocytes apparently routinely missed during manual screening of the FF. The results of a clinical trial in 19 patients are reported with one ongoing pregnancy from a “discarded” oocyte that was found by the device.

Overall Assessment

This study is novel and interesting and could be clinically significant if the device can be manufactured at reasonable cost and is shown to be equally effective by independent laboratories.

The approach to testing of this new device is sound and the study is well presented, although clarity can be improved and additional details may be needed.

Statistical methods are appropriate (but not central to the conclusions). References are appropriate.

That the device can reliably and automatically process FF and collect COCs while maintaining their viability has been clearly demonstrated. Although the authors' claim regarding the frequent and routine loss of some 50% of human oocytes during manual processing, independent of time pressures and COC numbers, is noteworthy and potentially impactful, this claim currently lacks sufficient explanation, detail, evidence and independent verification. Additional validation would help address important questions surrounding this unexpected observation.

Suggested improvements & clarifications:

This device appears to offer two advantages: automation of a labor-intensive procedure in the IVF laboratory and finding of eggs that may be missed during manual screening of FF. The authors seem to place more emphasis on the latter, even though many details that would support this finding are missing from the manuscript. The system combines oocyte collection with immediate denudation. This raises two concerns. First, it may make ICSI the default insemination method, at a time when credible evidence for a negative effect of ICSI on clinical outcomes in non-male factor patients with no indication for ICSI is accumulating. Second, the immediate removal of somatic supportive cells from oocytes after retrieval from the ovary may raise concerns regarding potential impact on viability and developmental potential. This concept has not been fully investigated - apart from some retrospective clinical studies providing weak evidence of its efficacy. Additional evidence or justification for this protocol would be beneficial here. Nonetheless, it is probably safe to conclude that enzymatic treatment during collection is the reason additional eggs are found by the device. Thus, the requirement for immediate denuding is likely a necessity and therefore can be considered a limitation of the design.

No limitations have been discussed and there is no description of potential problems/issues (encountered) during operation of the system. Were there none or is this an omission?

The authors should clarify what maximum volume of FF can be processed in one chip and whether more than one chip could be required for each case. Are these units disposable? How is sterility maintained within the system (e.g. hand-pouring of FF into a syringe rather than multiple fresh dishes) and were any instances of contamination noted? What about formation of blood clots during processing of bloody FF? Can the channels get blocked by blood clots or other debris?

The Abstract reports a clinical study with 500 patients. Elsewhere, the authors report 472 oocytes from 295 FF samples but the number of patients is not mentioned. In the supplementary information, the total number of patients is reported as 537. Please clarify the number of patients and clearly distinguish between the different studies: the screening of discarded FF (295 samples??) for recovery of eggs (with & without possibility of clinical use for the same patient); clinical study involving donated oocytes from 21 patients for “spiking” pre-screened autologous FF (with possibility of clinical use for the same patient), and the pilot trial involving the full use of the device to retrieve all oocytes in 19 patients. Were all studies conducted under the same IRB approval and consent? It is surprising that the consent form does not list loss or damage of oocytes as a potential risk. Instead, it states that “Risks associated with this study are not expected to be significant.” Can the authors explain?

Line 41; change suggested

Original: “with poor ovarian reserve due [to] low oocyte yield”

Suggested: “low oocyte yield resulting from poor ovarian reserve.”

Lines 45/46; change suggested

Original: “persistent gap between patient demand and IVF success rates”

Suggested: “persistent gap between patient demand and access to IVF”

Line 72; change suggested

Original: “automated device”

Suggested: “automated microfluidic-based device”

Line 99; change suggested

Original: “additional wash steps”

Suggested: “follicle flushing”

Line 112; Please specify gap sizes and how the sizes were determined considering the large variability in COC size.

Line 117; Please define inertial focusing briefly.

Lines 144 & 154; How is exposure time to hyaluronidase controlled (less than 120 seconds) and how is reaching “the desired oocyte denudation levels” determined? Is the latter through observation or is this automated based on size? Please clarify.

Line 160 & elsewhere; change suggested

Original: "clinical samples"

Suggested: "follicular fluid samples"

Please clarify if these samples were ever pooled or always used individually (maintaining patient identity).

Line 177; Please explain if the entire process in the device is monitored through the microscope?

Line 195; In the bovine experiments, the authors write that they used FF samples designated as “discard” by a trained embryologist after recovering “all” oocytes. The authors' use of quotation marks around 'discard' and 'all' conveys their uncertainty about the reliability of the manual process. Indeed, they are reporting that many human COCs were missed in discarded FF. Did the authors find any human oocytes in the samples that were spiked with bovine oocytes? If not, how could this be explained?

Line 310; Please clarify Gardner score 2EB.

Line 315 & elsewhere; change suggested

Original: "control group" / "clinical group" / "manual control group"

Suggested: "manually processed group," clearly distinguishing it as a comparison rather than a formal experimental control group.

Line 381; The authors write, "...portable aspect of the FIND-Chip will enable collection and preparation of eggs unconstrained by the availability of a trained workforce, or the need for an IVF laboratory." How will the eggs be managed after collection if not in a laboratory? Vitrification, ICSI, and culture all require a laboratory at this time. An individual without embryology skills may be able to decant the FF into the syringes for processing but handling the oocytes will still require embryology skills.

If I understand correctly, Figure 4 shows that a patient who had 9 eggs collected manually, had 11 additional eggs recovered by FIND-Chip. This significant discrepancy would typically prompt procedural reviews in clinical practice. A clearer explanation for such frequent discrepancies would help credibility. The number of eggs expected following ovarian stimulation is estimated based on the number of follicles seen on ultrasound as well as E2 levels at the time of hCG. Although these are often under-estimations particularly when many follicles develop and manual counts become less and less accurate (e.g., in patients who had 40-60+ eggs retrieved in this study (Fig S3)), the physician operator starts the operation with certain expectations and knowledge and is aware of how many follicles s/he has aspirated. Such details are not provided here. A case in which 20 eggs are expected and some 20 follicles are aspirated but only 9 eggs are found would be highly problematic for the embryology laboratory. How could this happen? One possibility is that physicians are in fact aspirating many more follicles than they "realize", including smaller follicles from which they do not expect to retrieve eggs. These COCs may be smaller and less conspicuous and perhaps easily missed during routine scanning. The data here show that nearly 60% of the FIND-Chip eggs are immature (MI & GV) or abnormal and this would be consistent with aspiration of smaller, immature/atretic? follicles.

Thus, oocyte aspiration protocols and approaches do come into play and should be considered when explaining the significant discrepancy between total number of eggs available and those found through manual screening, particularly considering the number of operators involved in this large network clinic. For this reason, it is also important that the authors specify whether the 3 local clinics that contributed material to the project were part of the same network or completely independent. If the former, this might be considered a limitation of the study and it would suggest that statements regarding the percentage of missed oocytes (about 50%) and the potential for significant improvements in IVF efficiency (17-50%) should be tempered. A proper study in an independent laboratory should include follicle counts/measurements obtained, preferably using newer automated technologies (e.g., Firtion et al., 2021) that can produce more accurate data and set realistic expectations with respect to the number of oocytes for each patient, and in turn make assessment of the success of oocyte retrieval in the lab more reliable.

Questionable terms & suggested alternatives:

Prognosis positive – Good prognosis

Purified denuded oocytes – Denuded oocytes

High purity oocytes – Isolated denuded oocytes

Clinical oocytes / clinical oocyte count / clinically recovered oocytes – Remove clinical

Developmental instrument - Prototype

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I am satisfied with the authors' responses to my comments.

Reviewer #2

(Remarks to the Author)

Thank you for the opportunity to review the revised version of this manuscript. The authors have responded constructively to the previous comments and have implemented several important improvements. These revisions strengthen the manuscript significantly and reflect the authors' effort to address reviewer concerns.

Key Results

This revised manuscript presents a comprehensive evaluation of the FIND Chip, a microfluidic platform designed to recover oocytes that may be missed during standard IVF follicular fluid (FF) processing. The study comprises four interconnected components:

1. Technology optimization with bovine samples (N = 10 experiments, 20 oocytes each), demonstrating high recovery and preservation of oocyte quality.
2. Validation using human oocytes (N = 21 patients, 62 oocytes), with 100% recovery of spiked oocytes and discovery of previously undetected "stealth" oocytes.
3. Multi-center analysis of discarded follicular fluid (N = 537), revealing additional oocytes in ~55% of cases.
4. Pilot clinical integration study (N = 19), where 12 mature recovered oocytes were utilized, resulting in 1 ongoing pregnancy.

Significance

The FIND Chip platform represents a novel and potentially transformative addition to ART workflows. If validated in larger trials, this tool could improve oocyte yield, especially in poor responders, and influence IVF practice globally. The manuscript introduces a fresh solution to a longstanding procedural blind spot in embryology labs. Its translational potential is high.

Data & Methodology

Data quality is strong. The revised supplemental information presents clear figures and updated tables with defined error bars, improved legends, and methodology transparency. Embryo grading details have been expanded, and blinded assessment is confirmed. The response letter and revised manuscript clearly describe how feedback has been incorporated, including breakdown of study components and reference to CONSORT-style flowcharts.

The revised manuscript demonstrates strong technological validation through spiking studies and observational data. The conclusions regarding oocyte recovery efficiency are robust, well supported by multiple datasets, and the multi-center approach enhances generalizability. The statistical presentation has been improved in this revision, and the authors have clarified device throughput, FF handling, and addressed operator variability.

Suggested Improvements (Mostly Addressed)

The authors have responded constructively to the previous comments and have implemented several important improvements, including:

1. Clearer delineation of the four sub-studies (technology optimization, validation, observational recovery study, and pilot clinical use).
2. The statistical presentation has been improved in this revision with addition of detailed sample sizes and improved statistical reporting.
3. The authors have clarified device throughput, FF handling, and addressed operator variability.
4. Clarification that modeled live birth rate (LBR) projections are not measured outcomes, with more cautious interpretation in the revised discussion.
5. Expanded description of embryo grading criteria and improved labeling of figures.

These revisions strengthen the manuscript significantly and reflect the authors' effort to address reviewer concerns.

However, several limitations remain:

1. The pilot clinical study, while promising, is small (N = 19 patients) with one reported ongoing pregnancy. The conclusions regarding clinical benefit, embryo competence, and LBR improvement remain preliminary.
2. Some procedural details—such as FF handling time, device throughput, and operator variability—would still benefit from further elaboration, particularly for clinical adoption.

Future research could include prospective trial designs evaluating clinical outcomes (e.g., cumulative live birth rate) and external validation across varied IVF center protocols. More granular data on embryo developmental competence (euploidy rates, implantation outcomes) would also be beneficial.

Clarity & Context

The manuscript is clearly written, logically structured, and adheres to Nature's editorial standards. The abstract now better reflects the study's scope. Key terminology (e.g., "clinical utility," "stealth oocytes") has been refined. IRB approvals, conflict of interest disclosures, and plans for data/code availability are appropriately described.

Conclusion

Despite these limitations, the manuscript presents a novel, well-executed technology with strong translational potential. It offers a compelling early step toward automated oocyte recovery in IVF and is likely to stimulate follow-up studies in the field. I would suggest the final version appropriately emphasizes the pilot nature of the clinical data and includes a clearly stated roadmap for future validation.

Reviewer #3

(Remarks to the Author)

I would like to thank the authors for carefully addressing most of my previous questions and revising their manuscript accordingly. I appreciate the clarification that the manuscript is referring to the percentage of patients gaining oocytes with FIND-Chip rather than the percentage of oocytes missed. I have 3 additional comments that the authors may wish to consider but from my point of view, no further responses are needed. This is an innovative and thought-provoking study that makes a valuable contribution to the field.

- The range of missed oocytes shown in Figure 4A is 0-100%. If I correctly understand that "unique samples" represent individual patients, at least 10 patients appear to have had 33-100% of their oocytes missed during manual collection. Presumably, if the full volume of follicular fluid had been processed for all patients, additional oocytes might have been recovered, potentially increasing the overall recovery of extra oocytes. This observation remains both unexpected and concerning.

- The title of the paper suggests that the purpose of this invention was to "improve oocyte yield". Is this accurate or was the purpose to automate the retrieval procedure and finding of extra oocytes was one result of the automation protocol?

- Regarding the following two statements in the revised version: "The current FIND-Chip device has a number of design features that may limit its use in clinics that perceive certain cases may be at risk" and "Contrary to published data showing no benefit of ICSI in non-male factor infertility cases, the majority of autologous fresh IVF cycles in North America now utilize ICSI (79%)."

I would suggest that the phrase "perceive certain cases may be at risk" could be interpreted as dismissive and it may benefit from a more neutral framing. Likewise, the term "contrary to" may be better replaced with "despite", since the persistence of ICSI overuse occurs despite evidence showing no improvement, or even potential harm, in non-male factor infertility.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source. The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to Reviewers' Comments on:

"Microfluidic chip to process follicular fluid improves oocyte yield for IVF patients" by Mutlu *et al*

New suggested title: **"Development and early clinical evaluation of a microfluidic system to improve oocyte yield from follicular fluid in IVF patients"**

We thank the reviewers for their thoughtful comments and suggestions to improve our manuscript. Here is our point-by-point response:

Reviewer #1 (Remarks to the Author):

Reviewer comment: *I read this study with great interest. The authors, for the first time, report a microfluidic device (FIND-Chip) that automates the isolation and denudation of oocytes from follicular fluid (FF) for IVF. In addition to validation using a bovine model, they take an important step forward by demonstrating the device's potential clinical utility, successfully identifying otherwise discarded oocytes in approximately half of the samples tested from ~500 patients. They further support this with a small pilot study showing that these recovered oocytes can contribute additional blastocysts to the embryo cohort.*

The manuscript is well written, and the experiments follow a clear and logical trajectory. I believe this technology represents a meaningful advance in IVF, with implications for improving success rates, workflow efficiency, and accessibility. I have only minor comments:

1) Could the authors discuss next steps for evaluating this technology beyond the current pilot? For instance, would a randomized controlled trial (RCT) comparing standard embryologist-led screening versus the FIND-Chip, perhaps using the FF from the same patient, be feasible or informative?

Authors' response: We thank the reviewer for their suggestion. We have added the following discussion on a future RCT in the manuscript on lines 335-345:

"To quantitatively assess the aggregate clinical impact of the FIND-Chip technology, randomized controlled trials could be designed either by randomization of patients performing IVF retrievals or splitting of follicular fluid samples of individual patients between automated and manual screening arms. For both approaches the primary endpoint would be the blastocyst utilization rate, i.e. number of fresh-transferred plus frozen blastocysts per retrieval. Secondary endpoints would include implantation rates per transfer and cumulative live birth rates per transfer. We estimate that the timeline for an adequately powered study would be approximately 1-2 years to collect the primary

objective but would be up to 2 to 3 years to collect the cumulative live birth outcomes, similar to some recent trials examining the use of sequential imaging and artificial intelligence (AI) to select embryos for transfer^{19,42}.”

Reviewer comment: 2) In IVF cycles triggered with a GnRH agonist, embryologists often visually confirm the presence of oocytes in the FF as an indirect check that the trigger worked. Do the authors envision incorporating an alert mechanism in the device to flag when an oocyte is identified? Alternatively, could the embryologist still briefly inspect the fluid before loading it into the device after the first egg is seen? A brief discussion on this could enhance the manuscript.

Authors’ response: We thank the reviewer for raising this point. We added the following discussion in the manuscript on lines 382-387 as suggested by the reviewer:

“Finally, the current design does not signal when the first oocyte is found in the follicular fluid, and the presence of oocytes is known at the completion of processing as opposed to confirmation of oocytes during the retrieval process. Future iterations of the system are envisioned to provide a means of oocyte identification during processing and presenting oocytes to enable conventional insemination, if required.”

Reviewer comment: 3) Line 40: Recommend replacing “prognosis positive” with “good prognosis.”; Line 345: The symbol for “less than or equal to 10” does not render properly, please correct the formatting; Lines 455–457: Please indicate whether a normality test was conducted prior to applying ANOVA.

Authors’ response: We edited line 40 as suggested and corrected line 345. We added the following section regarding the statistical methods:

“We used the Brown-Forsythe and Welch one-way ANOVA with Dunnett’s T3 multiple comparisons in analyzing fertilization and blastomere experiment results to compare all the groups. Normality was assessed with the Shapiro-Wilk test and was not rejected ($p > 0.05$).”

Reviewer #2 (Remarks to the Author):

Reviewer comment: *Thank you for the opportunity to review this manuscript, which presents the microfluidic FIND-Chip designed to recover previously undetected oocytes from follicular fluid in IVF.*

A. Summary of the Key Results

This manuscript reports findings from four distinct investigations using the FIND-Chip device:

- 1. Technology optimization using clinical samples spiked with bovine oocytes (N = 10 experiments with 20 oocytes each), demonstrating high isolation and denudation efficiency and equivalent or slightly improved embryo development rates.*
 - 2. Technology validation using donated human oocytes (N = 21 patients; 62 oocytes), showing 100% recovery of spiked-in oocytes and identification of additional “stealth” oocytes not detected by embryologists.*
 - 3. Multi-clinic investigation of discarded, pre-screened follicular fluid from IVF centers (N = 537 patients), where FIND-Chip identified extra oocytes in 54.9% of cases, regardless of the patient’s original oocyte yield.*
 - 4. Clinical utilization pilot study (N = 19 patients), where recovered oocytes were returned to the clinical workflow, resulting in one ongoing pregnancy.*
- These are compelling early findings that suggest the system’s promise for clinical use.*

B. Originality and Significance

The work is novel and impactful. Demonstrating that a substantial proportion of oocytes are missed during standard IVF workflows and can be recovered by microfluidic automation addresses a longstanding challenge in reproductive medicine. Increasing the usable oocyte pool—particularly for low responders—has potential to improve patient outcomes and reduce the number of cycles needed.

C. Data & Methodology: Validity, Quality, Presentation

The methodology is generally sound, with well-documented optimization in bovine and validation in human oocytes. The large, multi-clinic dataset strengthens observational findings. However:

- 1) Clinical investigations involving human oocytes are small, especially the pilot utilization study (N = 19, with only 12 mature extra oocytes used for ICSI). Conclusions about embryo competence and pregnancy potential should be framed as preliminary.*

Authors’ response: We agree. We revised lines 225-228 of the manuscript to clarify the preliminary nature of the results and the need for a larger cohort study for an adequately powered statistical comparison as follows:

“While a larger cohort is needed to assess the full impact of extra oocytes (via an adequately powered statistical comparison) recovered by FIND-Chip, preliminary results

show that FIND-Chip contributed additional usable embryos for patient treatment, increasing the likelihood of successful outcomes.”

Reviewer comment: 2) Define the “fragmented or deformed” oocyte category (13.4% of extras) and clarify whether these were viable. Indicate whether atypical morphology affected device performance (e.g., clogging or failure).

Authors’ response: We added lines 280-283 to clarify the “fragmented or deformed” oocyte category, and whether it affected device performance as follows:

“Fragmented or deformed oocytes refer to the eggs with morphological abnormalities that preclude ICSI due to cytoplasmic fragmentation, irregular oolemma/zona or dead. Processing of these atypical oocytes did not impair device operation.”

Reviewer comment: 3) Acknowledge that only 17.9% of FF samples were processed in full volume, potentially underestimating extra-oocyte prevalence. Explain how partial volumes were selected and whether sampling was representative.

Authors’ response: We reviewed our data analysis and corrected the fraction of samples that were processed in full volume to 31.8%, which was edited on lines 272-274 as follows:

“Of note, the full volume of the pre-screened FF was processed only for a fraction (31.8%) of the cases, primarily to allow testing of more samples from different patients (See SI for details).”

We added a discussion on FF volume in the supplementary information (SI) lines 90-101 as follows:

“We processed the full volume of the discarded FF for only 31.8% (185/582) of the patient samples. There were several considerations that determined whether we opted to fully process a sample, including sample volume, sample age (time between retrieval and process) and number of available samples. Typically, if there were other available samples to run, we processed only 55 mL from a sample to allow running FF samples from more patients (albeit without being able to run full volume) to evaluate device performance across a larger patient cohort. The 55 mL was selected based on the FF reservoir volume of the prototype system. The reservoir volume does not limit the amount of FF that the system can process, but requires the user to manually refill the reservoir and adds sample loading time to the process. Overall, in the general patient group, the average discarded FF volume collected was $V = 96.3\text{ mL}$, the average processed FF volume was $V = 56.9\text{ mL}$, and the maximum FF volume that was processed was $V = 200\text{ mL}$.”

Reviewer comment: 4) Discuss possible reasons for inter-clinic variability in extra-oocyte rates (Table 1), including procedural or patient population differences.

Authors' response: We added additional information on Table 1: "Heparin use during retrieval" and "Time allocated per retrieval", to highlight some differences between these clinics in terms of retrieval practices. We have added the following discussion in lines 300-305, as suggested:

Notably, the extra oocytes were recovered with the same consistency in samples from 4 independent clinics which had differences in standard clinical practices such as the use of anticoagulants during retrieval and time allocated for retrieval per patient. Based on the results shown in Table 1, we observed no significant variations in recovery of extra oocytes between clinics despite the differences in practice.

Reviewer comment: D. Appropriate Use of Statistics & Uncertainty Management

5) Statistical comparisons are appropriate but underpowered in the pilot trial. Variability in fertilization and blastocyst rates should be addressed, and p-values provided where available. Clarify whether observed differences (e.g., 63% vs. 49% blastocyst rate) were statistically significant or descriptive only. Ensure all error bars are defined and statistical assumptions stated. Future studies should consider more robust models (e.g., mixed-effects regression) to account for patient- and clinic-level variability.

Authors' response: We agree. We clarified that no statistical comparison was performed in the pilot trial and the comparison was descriptive only on lines 222-225 as follows:

"Grades of the embryos did not indicate any difference between the FIND-Chip and the manual screening group (Fig 3C). We did not perform statistical comparison between groups based on the low number of patients, and the plotted results are only descriptive.

We reviewed all the statistical comparisons in the paper to ensure the error bars, assumptions, p values, and other relevant metrics are stated. We have added a discussion on a future, multi-clinic RCT in the manuscript (lines 335-345):

"To quantitatively assess the aggregate clinical impact of the FIND-Chip technology, randomized controlled trials could be designed either by randomization of patients performing IVF retrievals or splitting of follicular fluid samples of individual patients between automated and manual screening arms. For both approaches the primary endpoint would be the blastocyst utilization rate, i.e. number of fresh-transferred plus frozen blastocysts per retrieval. Secondary endpoints would include implantation rates per transfer and cumulative live birth rates per transfer. We estimate that the timeline for an adequately powered study would be approximately 1-2 years to collect the primary objective but would be up to 2 to 3 years to collect the cumulative live birth outcomes,

similar to some recent trials examining the use of sequential imaging and artificial intelligence (AI) to select embryos for transfer^{19,42}.”

Reviewer comment: E. Conclusions: Robustness & Reliability

6) This is a strong and promising study introducing a potentially transformative technology in IVF. The technological performance is convincing, but clinical impact evidence is early-stage. The authors should clearly delineate the four studies, expand clinical validation in larger prospective trials, and address the above methodological and reporting points to maximize impact and suitability for Nature Medicine.

Authors’ response: We thank the reviewer for their comment. We acknowledge that the results regarding technology development and optimization can be better delineated from the early clinical validation results. Thus, we shortened the technology description and validation section of the main manuscript, and moved some details regarding device and process optimization to the supplementary information. We added an IRB table (Table 2) in the methods section to further outline and explain the different studies, and added a discussion on a future RCT in the manuscript lines 335-345, as shown above.

F. Suggested Improvements

Reviewer comment: 1. *The current title does not reflect the multi-phase nature of the work. Consider rephrasing to highlight both the technical development and the early clinical validation, e.g.: “Development and Early Clinical Evaluation of a Microfluidic System for Oocyte Recovery from Discarded Follicular Fluid”*

Authors’ response: We thank the reviewer for their feedback and appreciate the suggestion of an alternative title which has been incorporated.

Reviewer comment: 2. *Consider publishing the four distinct studies as stand-alone manuscripts for clarity and depth.*

Authors’ response: We thank the reviewer for their comment. We believe the presented technology content, pilot study, and multi-clinical demonstration collectively support the narrative that FIND-Chip system can be a transformative technology in clinical IVF practice, and suitable for the newly incorporated title of the paper as suggested by the reviewer. We also made edits in the manuscript in an effort to better streamline and clarify different studies, including an IRB table (Table 2) in the methods section.

Reviewer comment: 3. Provide a CONSORT-style diagram summarizing patient inclusion and clinical embryo outcomes.

Authors' response: We thank the reviewer for this suggestion. Based on the scope and non-limiting inclusion criteria of our studies, we were able to summarize the criteria and the clinical embryo outcomes in Table 2, which is included in the methods section of the manuscript.

Reviewer comment: 4. Discuss the timeline, design, and endpoints for a prospective cohort or randomized trial to evaluate FIND-Chip's clinical impact (e.g., cumulative live birth rate).

Authors' response: We thank the reviewer for their suggestion. We have added a discussion on a future RCT in the manuscript lines 335-345 as follows:

“To quantitatively assess the aggregate clinical impact of the FIND-Chip technology, randomized controlled trials could be designed either by randomization of patients performing IVF retrievals or splitting of follicular fluid samples of individual patients between automated and manual screening arms. For both approaches the primary endpoint would be the blastocyst utilization rate, *i.e.* number of fresh-transferred plus frozen blastocysts per retrieval. Secondary endpoints would include implantation rates per transfer and cumulative live birth rates per transfer. We estimate that the timeline for an adequately powered study would be approximately 1-2 years to collect the primary objective but would be up to 2 to 3 years to collect the cumulative live birth outcomes, similar to some recent trials examining the use of sequential imaging and artificial intelligence (AI) to select embryos for transfer^{19,42}.”

Reviewer comment: 5. Address FF sample handling time, device throughput, and operator variability.

Authors' response: We added these details to the main manuscript on lines 430-431:

“Typically, it took about one minute of handling time for the operator to aspirate the FF from the pot, load it into the reservoir and start the process.”

441-443:

“Overall, operator interaction with the instrument was limited to filling of the three reservoirs with corresponding fluids (MHM-C, HYA and FF), which did not cause variability in experimental outcomes.”

and supplementary information lines 70-74:

“Under these optimized parameters, the sample processing time (to released product) using FIND-Chip was approximately 25-30 minutes, after the sample is fully loaded into the system. Sample loading time scaled linearly with the sample volume and was approximately 5 minutes for a 50mL FF sample.”

Reviewer comment: 6. *Ensure all error bars and statistical comparisons are clearly defined.*

Authors’ response: We reviewed the graphs and ensured all the relevant details are provided where applicable.

Reviewer comment: 7. *Include embryo grading methodology and whether assessments were blinded.*

Authors’ response: We added these details as suggested on lines 215-219 as follows:

“The quality of blastocysts was assessed on days 5, 6 and 7 (per Gardner criteria³³) per the clinic’s standard protocol by the embryologist on duty who was not part of the research team. The embryologists were not blinded to whether embryos came from manual screening or FIND-chip because FIND-Chip oocytes/embryos were tracked separately as required by the IRB protocol.”

Reviewer comment: 8. *Clearly distinguish modeled projections of live birth rate improvement from measured outcomes. In discussion section (lines 359-366), the study’s projections of improved live birth rates are based on models, not measured clinical outcomes. Actual improvements remain unproven and should be presented as potential benefits pending further clinical validation.*

Authors’ response: We agree. We removed the discussion section with the simplified model of live birth rate per oocyte and projections. Instead, we added the clinical data from a retrospective, large-cohort (16,474 patients) study that demonstrates the clear correlation between clinically observed cumulative live birth rate (CLBR) vs. number of oocytes collected from the index cycle (Fig 4B). This data is from a published study and we obtained the raw data by requesting it from the authors of the original study (Huttler et al, Achieving two live births from one ovarian stimulation cycle: the one-and-done approach revisited. Fertil. Steril. 122, 2025). On this plot, we also added dashed lines showing how the graph shifts with addition of extra oocytes retrieved per clinical oocyte count. We added the following section to lines 316-326:

“The strong positive correlation between the oocyte yield and clinically observed cumulative live birth rate (CLBR) is well established by multiple retrospective studies across different clinics with large patient cohorts ^{40,41}. In a recent report by Huttler et al.

where 16,474 patients were tracked, the clinical outcome of “one or more live births” using the oocytes from the index retrieval cycle was reported as: 14% (167/1,227) in patients who had 1-3 oocytes, 34% (1,672/4,933) in patients who had 4-9 oocytes, 53% (2,077/3,936) in patients who had 10-14 oocytes, 65% (3,091/4,729) in patients who had 15-25 oocytes, and 74% (1,216/1,649) in patients who had > 25 oocytes. Using the raw data from this study (provided by request from the authors), we plotted the increase in the CLBR per oocyte, and how the corresponding CLBR would shift upward when extra oocytes are available for the patients (Fig 4B)⁴¹.”

Reviewer comment: 9. Correct “pre-implementation genetic testing” to “pre-implantation genetic testing for aneuploidy (PGT-A).”

Authors’ response: We corrected the text as suggested.

Reviewer #3 (Remarks to the Author):

Reviewer comment:

SUMMARY

This study provides proof of concept for a microfluidic device (FIND-Chip) that automates isolation and denudation of cumulus-oocyte-complexes (COCs) from follicular fluid (FF) aspirates obtained for IVF. Under experimental conditions, the device is shown to efficiently isolate and denude bovine and human COCs. Furthermore, re-screening of pre-screened FF shows that the device identifies additional oocytes apparently routinely missed during manual screening of the FF. The results of a clinical trial in 19 patients are reported with one ongoing pregnancy from a “discarded” oocyte that was found by the device.

Overall Assessment

This study is novel and interesting and could be clinically significant if the device can be manufactured at reasonable cost and is shown to be equally effective by independent laboratories.

The approach to testing of this new device is sound and the study is well presented, although clarity can be improved and additional details may be needed.

Statistical methods are appropriate (but not central to the conclusions). References are appropriate.

That the device can reliably and automatically process FF and collect COCs while maintaining their viability has been clearly demonstrated. Although the authors' claim regarding the frequent and routine loss of some 50% of human oocytes during manual processing, independent of time pressures and COC numbers, is noteworthy and potentially impactful, this claim currently lacks sufficient explanation, detail, evidence and independent verification. Additional validation would help address important questions surrounding this unexpected observation.

Authors' response: We would like to clarify that our data shows that “in greater than 50% of the cases at least one oocyte is missed during manual screening”, and not that 50% of total oocytes are missed. To provide further evidence to this phenomenon, we added new data from a fourth independent clinic “Clinic D”. We also detailed key differences in retrieval practices between these independent clinics in Table 1 to highlight that the system performance is mostly agnostic to these differences.

Reviewer comment: *This device appears to offer two advantages: automation of a labor-intensive procedure in the IVF laboratory and finding of eggs that may be missed during manual screening of FF. The authors seem to place more emphasis on the latter, even*

though many details that would support this finding are missing from the manuscript. The system combines oocyte collection with immediate denudation. This raises two concerns. First, it may make ICSI the default insemination method, at a time when credible evidence for a negative effect of ICSI on clinical outcomes in non-male factor patients with no indication for ICSI is accumulating. Second, the immediate removal of somatic supportive cells from oocytes after retrieval from the ovary may raise concerns regarding potential impact on viability and developmental potential. This concept has not been fully investigated - apart from some retrospective clinical studies providing weak evidence of its efficacy. Additional evidence or justification for this protocol would be beneficial here. Nonetheless, it is probably safe to conclude that enzymatic treatment during collection is the reason additional eggs are found by the device. Thus, the requirement for immediate denuding is likely a necessity and therefore can be considered a limitation of the design.

Authors' response: We thank the reviewer for their suggestion and acknowledge that these limitations: "ICSI being the default insemination method" and "Immediate denudation of oocytes" warrant a discussion in the manuscript. As suggested, we added the following discussion on these topics in the manuscript lines 368-387:

"The current FIND-Chip device has a number of design features that may limit its use in clinics that perceive certain cases may be at risk. Firstly, the FIND-Chip system delivers denuded oocytes, making it necessary to perform ICSI as the default insemination method. This decision was driven by the fact that contrary to published data showing no benefit of ICSI in non-male factor infertility cases^{50,51} the majority of autologous fresh IVF cycles in North America now utilize ICSI (79%)⁵². Secondly, cumulus oocytes are denuded immediately instead of a "wait time" after oocyte removal from FF. Interestingly, this practice seems to have ensued more so because of logistical reasons, as highlighted in a recent study by Barcena *et al.*⁵³ who concluded that the effective window of time for insemination by ICSI might be wider than previously thought. Even though there are differences among clinical practices in terms of a "wait time" between removal of oocytes from FF and denudation, recent findings found no impact of minimizing this wait time on ICSI fertilization rates per MII injected, blastocyst rates and live birth rates⁵⁴. Interestingly, it is also common practice to denude oocytes in the first hour prior to vitrification⁵⁵ and egg banks in the United States recommend immediate denudation after collection in their own protocols. Finally, the current design does not signal when the first oocyte is found in the follicular fluid, and the presence of oocytes is known at the completion of processing as opposed to confirmation of oocytes during the retrieval process. Future iterations of the system are envisioned to provide a means of presenting oocytes to enable conventional insemination and oocyte identification during processing, if required."

Reviewer comment: *No limitations have been discussed and there is no description of*

potential problems/issues (encountered) during operation of the system. Were there none or is this an omission?

Authors' response: We thank the reviewer for their suggestion. We added a discussion on system limitations in response to the previous comment (lines 368-387), as shown above. We also detailed the issues we encountered during the operation of the prototype system on lines 443-446 as follows:

“We had failed processes very infrequently, which were due to errors in chip assembly (e.g. tubing not properly connected to chip) or manufacturing process (e.g. bad bonding of PDMS to glass), which are expected due to the early prototype stage of the kit and the instrument.”

Reviewer comment: *The authors should clarify what maximum volume of FF can be processed in one chip and whether more than one chip could be required for each case. Are these units disposable? How is sterility maintained within the system (e.g. hand-pouring of FF into a syringe rather than multiple fresh dishes) and were any instances of contamination noted? What about formation of blood clots during processing of bloody FF? Can the channels get blocked by blood clots or other debris?*

Authors' response: We thank the reviewer for their comments and suggestions. There is no maximum, limiting volume that could be processed by a chip. In this study, the maximum FF volume we processed was 200 mL. We added details and a discussion on FF volume in supplementary information lines 89-101 as follows:

“We processed the full volume of the discarded FF for only 31.8% (185/582) of the patient samples. There were several considerations that determined whether we opted to fully process a sample, including sample volume, sample age (time between retrieval and process) and number of available samples. Typically, if there were other available samples to run, we processed only 55 mL from a sample to allow running FF samples from more patients (albeit without being able to run full volume) to evaluate device performance across a larger patient cohort. The 55 mL was selected based on the FF reservoir volume of the prototype system. The reservoir volume does not limit the amount of FF that the system can process, but requires the user to manually refill the reservoir and adds sample loading time to the process. Overall, in the general patient group, the average discarded FF volume collected was $V = 96.3\text{ mL}$, the average processed FF volume was $V = 56.9\text{ mL}$, and the maximum FF volume that was processed was $V = 200\text{ mL}$.”

We added details regarding disposable chips and sterility on lines 409-414 as follows:

“PDMS chips are disposable, but not suitable for large scale clinical use because of limitations in scalable manufacturing. Injection molded plastic chips are disposable, can be inexpensively fabricated at scale, and suitable (along with disposable, single-use tubing

and reservoirs) for various sterilization methods (e.g. E-beam, gamma irradiation). Once the FF is loaded into the system reservoir, the fluidic system is closed, and tissue only interacts with sterilized, disposable components.”

We added details regarding limitations around running bloody FF samples as suggested on lines 424-430:

“Two of the participating clinics used an anti-coagulant (heparin) during FF retrieval, while the other two did not. At the time of processing, occasionally blood clots were observed in the pre-screened FF samples. Small blood clots in the sample were trapped by the filter module of the chip during processing, and did not impact device process. On occasion, a multi-millimeter scale blood clot was observed in the FF sample, which was not aspirated by the operator.”

Reviewer comment: *The Abstract reports a clinical study with 500 patients. Elsewhere, the authors report 472 oocytes from 295 FF samples but the number of patients is not mentioned. In the supplementary information, the total number of patients is reported as 537. Please clarify the number of patients and clearly distinguish between the different studies: the screening of discarded FF (295 samples??) for recovery of eggs (with & without possibility of clinical use for the same patient); clinical study involving donated oocytes from 21 patients for “spiking” pre-screened autologous FF (with possibility of clinical use for the same patient), and the pilot trial involving the full use of the device to retrieve all oocytes in 19 patients.*

Were all studies conducted under the same IRB approval and consent? It is surprising that the consent form does not list loss or damage of oocytes as a potential risk. Instead, it states that “Risks associated with this study are not expected to be significant.” Can the authors explain?

Authors’ response: We thank the reviewer for pointing this out. We corrected the abstract with the exact number of patients, which is 582 with the addition of data from a fourth independent clinic. We edited the introduction section (lines 91-94) to further clarify the number of patients (earlier version cited 295, which was the number of patients where at least one extra oocyte was found) as follows:

“FIND-chip recovered a total of 582 oocytes from screened and discarded FF samples of 582 patients, indicating the potential harm to patient treatment -particularly for those with low oocyte yield- as a result of the shortcomings of current standard of care.”

We also expanded Table 1 to provide additional information on these samples, and re-clarified these numbers where applicable.

There were different IRB approvals required for different sections of this study. Only for one of the IRBs, the oocytes are returned back to clinical workflow. Since these oocytes are

recovered from the already screened and discarded follicular fluid, damage to or even loss of oocytes are considered not to increase the risk for patient's treatment. These IRB requirements and the associated sections are added as **Table 2** in the methods section.

Reviewer comment: Line 41; change suggested

Original: "with poor ovarian reserve due [to] low oocyte yield"

Suggested: "low oocyte yield resulting from poor ovarian reserve."

Authors' response: We edited the text as suggested.

Reviewer comment: Lines 45/46; change suggested

Original: "persistent gap between patient demand and IVF success rates"

Suggested: "persistent gap between patient demand and access to IVF"

Authors' response: We edited the text as suggested.

Reviewer comment: Line 72; change suggested

Original: "automated device"

Suggested: "automated microfluidic-based device"

Authors' response: We edited the text as suggested.

Reviewer comment: Line 99; change suggested

Original: "additional wash steps"

Suggested: "follicle flushing"

Authors' response: We edited the text as suggested.

Reviewer comment: Line 112; Please specify gap sizes and how the sizes were determined considering the large variability in COC size.

Authors' response: We added the following details to the manuscript on lines 120-124 as the reviewer suggested:

"The filtration size cut-off starts at 3 mm gap size between posts, and gradually reduces to a final gap size of 169 μm . These gap sizes were selected to allow trapping of COCs during sample loading at various filter stages, and eventual release of oocytes after hyaluronidase treatment when they are stripped from the cumulus tissue and only retain their tight corona layer."

Reviewer comment: Line 117; Please define inertial focusing briefly.

Authors' response: We added a brief description of inertial focusing as suggested on line 127.

Reviewer comment: Lines 144 & 154; How is exposure time to hyaluronidase controlled (less than 120 seconds) and how is reaching “the desired oocyte denudation levels” determined? Is the latter through observation or is this automated based on size? Please clarify.

Authors’ response: We thank the reviewer for the comment. This section where technology was further described was moved to the supplementary information while re-organizing the paper. We added more details on lines SI 23-26 to further explain how exposure time is controlled as follows:

“Importantly, HYA treatment is immediately followed by a wash step to replace the HYA solution interacting with the tissue with media (MHM-C), which effectively stops the exposure of oocytes to HYA and limits the HYA exposure time to less than 120 seconds, in line with clinical standards.”

We edited the text to clarify “the desired oocyte denudation levels” on lines SI 35-38 as follows:

The denudation and wash cycles are repeated until the oocytes are fully denuded, and the desired product enrichment levels are achieved. The number of denudation and wash cycles are pre-set based on experimental observations during technology optimization studies.

Reviewer comment: Line 160 & elsewhere; change suggested
Original: “clinical samples”
Suggested: “follicular fluid samples”

Authors’ response: We thank the reviewer for the comment. We edited the line as suggested and at other places where applicable.

Reviewer comment: Please clarify if these samples were ever pooled or always used individually (maintaining patient identity).

Authors’ response: We added these details as suggested on lines 431-432 as follows:

Samples from different patients were never pooled (mixed) and only one patient’s sample was processed at a time.

Reviewer comment: Line 177; Please explain if the entire process in the device is monitored through the microscope?

Authors’ response: We added these details as suggested on lines 451-454 as follows:

In the majority of runs, the device was not placed under a microscope to monitor the process. We utilized the microscope in a subset of runs (mostly in early development

stages) for characterization and to acquire images.

Reviewer comment: Line 195; In the bovine experiments, the authors write that they used FF samples designated as ““discard” by a trained embryologist after recovering “all” oocytes” The authors' use of quotation marks around 'discard' and 'all' conveys their uncertainty about the reliability of the manual process. Indeed, they are reporting that many human COCs were missed in discarded FF. Did the authors find any human oocytes in the samples that were spiked with bovine oocytes? If not, how could this be explained?

Authors' response: The reviewer is correct that human oocytes were routinely recovered in the samples that were spiked with bovine oocytes as well. This is explained in the paper on lines 191-193 as shown below, and we edited this section to add an image of a human oocyte recovered along with bovine samples as a supplementary figure:

“Of note, this outcome was consistent with our observations during bovine oocyte spiking experiments, where we had frequently identified human oocytes isolated from the pre-screened FF samples among bovine oocytes (FIG S6).”

Reviewer comment: Line 310; Please clarify Gardner score 2EB.

Authors' response: We clarified that we mean Early Blastocyst grade and above by 2EB+. We edited this detail on line 214.

Reviewer comment: Line 315 & elsewhere; change suggested
Original: “control group” / “clinical group” / “manual control group”
Suggested: “manually processed group,” clearly distinguishing it as a comparison rather than a formal experimental control group.

Authors' response: We agree. We changed the wording from “control” to “manual screening” where applicable.

Reviewer comment: Line 381; The authors write, “...portable aspect of the FIND-Chip will enable collection and preparation of eggs unconstrained by the availability of a trained workforce, or the need for an IVF laboratory.” How will the eggs be managed after collection if not in a laboratory? Vitrification, ICSI, and culture all require a laboratory at this time. An individual without embryology skills may be able to decant the FF into the syringes for processing but handling the oocytes will still require embryology skills.

Authors' response: We agree. We clarified this point and further elaborated about the “satellite clinic” concept on lines 357-366 as follows:

“The automated closed system, and portable aspect of the FIND-Chip will enable collection and preparation of eggs unconstrained by the availability of a trained workforce, or the need for an IVF laboratory. Together with other advances in miniaturized automation of vitrification, ICSI and embryo culture, this opens up the possibility of establishing and

operating smaller and lower-cost clinics to increase patient access to fertility treatment. For example, egg retrieval and oocyte vitrification could be completed at remote “satellite” clinics operated by larger IVF centers, which would drastically reduce the travel burden for the patient. Vitrified oocytes could be transported and handled at a central IVF lab for completion of the IVF procedure. In this scenario, FIND-Chip performing the isolation and denudation steps simultaneously leaves only the task of freezing of oocytes (and sperm), which could be completed either by an embryologist or an automated vitrification system.”

Reviewer comment: *If I understand correctly, Figure 4 shows that a patient who had 9 eggs collected manually, had 11 additional eggs recovered by FIND-Chip. This significant discrepancy would typically prompt procedural reviews in clinical practice. A clearer explanation for such frequent discrepancies would help credibility. The number of eggs expected following ovarian stimulation is estimated based on the number of follicles seen on ultrasound as well as E2 levels at the time of hCG. Although these are often under-estimations particularly when many follicles develop and manual counts become less and less accurate (e.g., in patients who had 40-60+ eggs retrieved in this study (Fig S3)), the physician operator starts the operation with certain expectations and knowledge and is aware of how many follicles s/he has aspirated. Such details are not provided here. A case in which 20 eggs are expected and some 20 follicles are aspirated but only 9 eggs are found would be highly problematic for the embryology laboratory. How could this happen? One possibility is that physicians are in fact aspirating many more follicles than they "realize", including smaller follicles from which they do not expect to retrieve eggs. These COCs may be smaller and less conspicuous and perhaps easily missed during routine scanning. The data here show that nearly 60% of the FIND-Chip eggs are immature (MI & GV) or abnormal and this would be consistent with aspiration of smaller, immature/atretic? follicles.*

Thus, oocyte aspiration protocols and approaches do come into play and should be considered when explaining the significant discrepancy between total number of eggs available and those found through manual screening, particularly considering the number of operators involved in this large network clinic. For this reason, it is also important that the authors specify whether the 3 local clinics that contributed material to the project were part of the same network or completely independent. If the former, this might be considered a limitation of the study and it would suggest that statements regarding the percentage of missed oocytes (about 50%) and the potential for significant improvements in IVF efficiency (17-50%) should be tempered. A proper study in an independent laboratory should include follicle counts/measurements obtained, preferably using newer automated technologies (e.g., Firtion et al., 2021) that can produce more accurate data and set realistic expectations with respect to the number of oocytes for each patient, and in turn make assessment of the success of oocyte retrieval in the lab more reliable.

Authors' response: The reviewer is correct that in one case, the system recovered 11 eggs from a patient who had 9 eggs collected in manual screening, and this particular case was communicated to the clinic. In the revised manuscript, within the added data from a new independent clinic, there has been a similar case where the clinical count was 23 and we

recovered 10 extra oocytes. We believe that these cases, and few other similar ones, are outliers given that most of our extra oocytes are within the 1 to 4 range. While the level of correlation between the number of follicles vs. the number of recovered oocytes has been an ongoing topic of debate to this date (for example Bosdou et al, *Can we estimate the minimum number of oocytes retrieved based on follicle diameter on trigger day? A cohort study, 2025*), we trust that the technology presented in this paper may aid in understanding this correlation better.

In the revised submission, addition of data from another clinic (“Clinic D”) brings the total number of independent clinics to 4, number of clinicians to 39, and number of embryologists to 61 who worked on the samples we processed. These clinics are not part of the same network, and have differences in their clinical practices. We added this extra information to Table 1, in terms of “heparin use during retrieval” and “time allocated per retrieval”. We agree that aspiration of smaller/immature follicles may contribute to having smaller COCs, increasing the likelihood of missing them during manual screening. However, consistent recovery of mature M2 oocytes across 4 clinics, which constitutes >40% of the extra oocyte population suggests there are several factors -independent of clinical practices- that contribute to these missed oocytes.

Reviewer comment: *Questionable terms & suggested alternatives:*

Prognosis positive – Good prognosis

Purified denuded oocytes – Denuded oocytes

High purity oocytes – Isolated denuded oocytes

Clinical oocytes / clinical oocyte count / clinically recovered oocytes – Remove clinical

Developmental instrument – Prototype

Authors’ response: We agree. We changed the wording as suggested by the reviewer.

Response to Reviewers' Comments on:

"Development and early clinical evaluation of a microfluidic system to improve oocyte yield from follicular fluid in IVF patients" by Mutlu *et al*

We thank the reviewers for their thoughtful comments and suggestions to improve our manuscript. Here is our point-by-point response:

Reviewer #1:

Reviewer comment: *I am satisfied with the authors' responses to my comments.*

Authors response: We thank the reviewer for their comments and suggestions which significantly improved the paper.

Reviewer #2:

Reviewer comment: *Thank you for the opportunity to review the revised version of this manuscript. The authors have responded constructively to the previous comments and have implemented several important improvements. These revisions strengthen the manuscript significantly and reflect the authors' effort to address reviewer concerns.*

Key Results

This revised manuscript presents a comprehensive evaluation of the FIND Chip, a microfluidic platform designed to recover oocytes that may be missed during standard IVF follicular fluid (FF) processing. The study comprises four interconnected components:

- 1. Technology optimization with bovine samples (N = 10 experiments, 20 oocytes each), demonstrating high recovery and preservation of oocyte quality.*
- 2. Validation using human oocytes (N = 21 patients, 62 oocytes), with 100% recovery of spiked oocytes and discovery of previously undetected "stealth" oocytes.*
- 3. Multi-center analysis of discarded follicular fluid (N = 537), revealing additional oocytes in ~55% of cases.*
- 4. Pilot clinical integration study (N = 19), where 12 mature recovered oocytes were utilized, resulting in 1 ongoing pregnancy.*

Significance

The FIND Chip platform represents a novel and potentially transformative addition to ART workflows. If validated in larger trials, this tool could improve oocyte yield, especially in poor responders, and influence IVF practice globally. The manuscript introduces a fresh solution to a longstanding procedural blind spot in embryology labs. Its translational potential is high.

Data & Methodology

Data quality is strong. The revised supplemental information presents clear figures and updated tables with defined error bars, improved legends, and methodology transparency. Embryo grading details have been expanded, and blinded assessment is confirmed. The response letter and revised manuscript clearly describe how feedback has been incorporated, including breakdown of study components and reference to CONSORT-style flowcharts.

The revised manuscript demonstrates strong technological validation through spiking studies and observational data. The conclusions regarding oocyte recovery efficiency are robust, well supported by multiple datasets, and the multi-center approach enhances generalizability. The statistical presentation has been improved in this revision, and the authors have clarified device throughput, FF handling, and addressed operator variability.

Suggested Improvements (Mostly Addressed)

The authors have responded constructively to the previous comments and have implemented several important improvements, including:

- 1. Clearer delineation of the four sub-studies (technology optimization, validation, observational recovery study, and pilot clinical use).*
- 2. The statistical presentation has been improved in this revision with addition of detailed sample sizes and improved statistical reporting.*
- 3. The authors have clarified device throughput, FF handling, and addressed operator variability.*
- 4. Clarification that modeled live birth rate (LBR) projections are not measured outcomes, with more cautious interpretation in the revised discussion.*
- 5. Expanded description of embryo grading criteria and improved labeling of figures.*

These revisions strengthen the manuscript significantly and reflect the authors' effort to address reviewer concerns. However, several limitations remain:

- 1. The pilot clinical study, while promising, is small (N = 19 patients) with one reported ongoing pregnancy. The conclusions regarding clinical benefit, embryo competence, and LBR improvement remain preliminary.*
- 2. Some procedural details—such as FF handling time, device throughput, and operator variability—would still benefit from further elaboration, particularly for clinical adoption.*

Author's response: We agree. Usability prototypes of this technology are under development and will be utilized for more exhaustive testing at clinical sites by laboratory personnel. This effort will increase the amount of both clinical data (including potential pregnancy outcomes) and usability data (FF handling time of the user, operator variability, etc.) to evaluate clinical adoption. Outcomes of that study, when completed, will be reported in a subsequent publication.

Future research could include prospective trial designs evaluating clinical outcomes (e.g.,

cumulative live birth rate) and external validation across varied IVF center protocols. More granular data on embryo developmental competence (euploidy rates, implantation outcomes) would also be beneficial.

Clarity & Context

The manuscript is clearly written, logically structured, and adheres to Nature's editorial standards. The abstract now better reflects the study's scope. Key terminology (e.g., "clinical utility," "stealth oocytes") has been refined. IRB approvals, conflict of interest disclosures, and plans for data/code availability are appropriately described.

Conclusion

Despite these limitations, the manuscript presents a novel, well-executed technology with strong translational potential. It offers a compelling early step toward automated oocyte recovery in IVF and is likely to stimulate follow-up studies in the field. I would suggest the final version appropriately emphasizes the pilot nature of the clinical data and includes a clearly stated roadmap for future validation.

Author's response: We thank the reviewer for their constructive comments. We have further reviewed and highlighted the pilot nature of the clinical data, where applicable. We believe that the roadmap includes a larger clinical study for further validation of FIND-Chip, which is described in the third paragraph of the Discussion section.

Reviewer #3:

Reviewer comment: *I would like to thank the authors for carefully addressing most of my previous questions and revising their manuscript accordingly. I appreciate the clarification that the manuscript is referring to the percentage of patients gaining oocytes with FIND-Chip rather than the percentage of oocytes missed. I have 3 additional comments that the authors may wish to consider but from my point of view, no further responses are needed. This is an innovative and thought-provoking study that makes a valuable contribution to the field.*

Author's response: We thank the author for their constructive feedback. We have provided our responses below to reviewer's final comments.

- *The range of missed oocytes shown in Figure 4A is 0-100%. If I correctly understand that "unique samples" represent individual patients, at least 10 patients appear to have had 33-100% of their oocytes missed during manual collection. Presumably, if the full volume of follicular fluid had been processed for all patients, additional oocytes might have been recovered, potentially increasing the overall recovery of extra oocytes. This observation remains both unexpected and concerning.*

Author's response: We agree. The reviewer has correctly understood the implications of Fig.4 - there are a significant number of patients where a substantial portion of oocytes are missed. This was also unexpected for the authors, and the information is shared with the

clinical sites.

- The title of the paper suggests that the purpose of this invention was to "improve oocyte yield". Is this accurate or was the purpose to automate the retrieval procedure and finding of extra oocytes was one result of the automation protocol?

Author's response: We thank the reviewer for raising this point. We have revised the manuscript title as follows:

"Development and clinical evaluation of a microfluidic device to automate and improve oocyte recovery from follicular fluid in patients undergoing in vitro fertilization"

- Regarding the following two statements in the revised version: "The current FIND-Chip device has a number of design features that may limit its use in clinics that perceive certain cases may be at risk" and "Contrary to published data showing no benefit of ICSI in non-male factor infertility cases, the majority of autologous fresh IVF cycles in North America now utilize ICSI (79%)."

I would suggest that the phrase "perceive certain cases may be at risk" could be interpreted as dismissive and it may benefit from a more neutral framing. Likewise, the term "contrary to" may be better replaced with "despite", since the persistence of ICSI overuse occurs despite evidence showing no improvement, or even potential harm, in non-male factor infertility.

Author's response: We made the following changes, per the reviewer's suggestion:

"The current FIND-Chip device has a few design features that may limit its use in certain clinics based on current practices"

"...Despite published data showing no benefit of ICSI in non-male factor infertility cases, the majority of autologous fresh IVF cycles in North America now utilize ICSI (79%)."