

INTRODUCTION

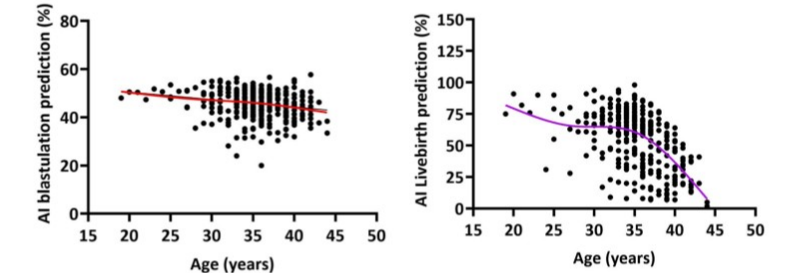
Determining oocyte quality is one of the main challenges in assisted reproductive technology since quality assessment have not much evolved beyond standard morphological observation. A deep learning model using artificial intelligence (AI), developed in Toronto, analyzes static images of oocytes images and has demonstrated promising performance in predicting their competence to develop into blastocyst.

AIM

This study aims to evaluate the variability of oocyte quality across different age groups as assessed by AI, under the hypothesis that variability increases with women's age. A secondary objective is to determine whether other variables such as ovarian reserve, body mass index, smoking, endometriosis severity, and type of ART treatment contribute to this variability.

METHOD

- Unicentric retrospective study at a private fertility clinic
- Elective fertility preservation, 326 cycles with only the first cycle considered per patients
- From March 2021 to December 2024
- All data regarding patients characteristics, fertility history, hormonal stimulation and treatment outcomes were collected
- Statistical comparisons between age groups and correlation analysis were performed following normality testing
- Appropriate parametric or non-parametric tests were used depending on data distribution
- GraphPad Prism 10.4.2, p-value < 0.05 as statistically significant



- Negative correlation between AI-blastulation and AI-livebirth prediction rates with age
- To assess dispersion across age, patients were stratified into four groups: <35, 35–37, 38–40, and ≥41 years.

RESULTS

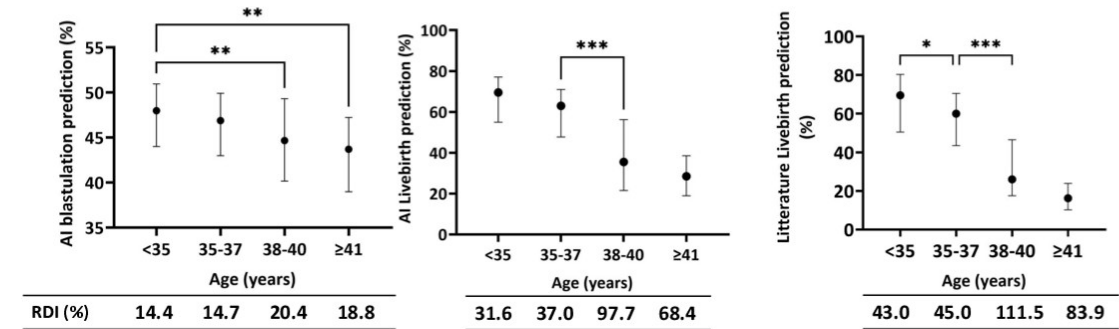
Patients and treatment characteristics

- No statistical differences in BMI, weight, smoking/vaping used, diagnosis of endometriosis stade III or IV, gravidity, parity, miscarriages for patients of each age group

	<35 n=122	35-37 n=126	38-40 n=52	≥41 n=26	P-value
Total dose of FSH (UI)	4361 ± 1581 4500 (2947, 5745)	4594 ± 1508 4800 (3300, 5760)	5466 ± 1144 5400 (4800, 6180)	5372 ± 1017 5295 (4680, 6240)	<35 vs 38-40/≥41 *** & * 35-37 vs 38-40 **
Stimulation duration (days)	11.5 ± 1.9 11.0 (10.0, 12.0)	11.5 ± 1.6 11.0 (10.0, 13.0)	12.1 ± 1.6 12.0 (11.0, 13.0)	11.7 ± 1.9 11.5 (10.7, 13.0)	<35 vs. 38-40 * NS
Protocols					
Antagonist	105 (86.1%)	108 (85.7%)	47 (90.4%)	24 (92.3%)	
PPOS	17 (13.9%)	18 (14.3%)	5 (9.6%)	2 (7.7%)	
Retrieved oocytes (n)	17.4 ± 10.2 15.5 (10.7, 22.2)	16.1 ± 9.7 14.0 (9.0, 20.0)	12.5 ± 8.8 10.0 (5.0, 17.7)	10.3 ± 5.7 9.0 (5.0, 13.5)	<35 vs 38-40/≥41 ** 35-37 vs ≥41 *
Number of MII (n)	13.0 ± 8.0 12.0 (7.0, 17.0)	11.7 ± 7.2 11.0 (6.7, 15.0)	9.1 ± 6.9 6.5 (4.0, 14.7)	7.7 ± 4.5 7.0 (4.0, 10.5)	<35 vs 38-40 /≥41 ** <35 vs ≥41 *
Evaluated oocytes (n)	12.8 ± 8.0 12.0 (7.0, 16.2)	11.9 ± 8.1 11.0 (7.0, 15.0)	9.0 ± 6.9 6.5 (4.0, 13.7)	7.6 ± 4.5 7.0 (4.0, 10.5)	<35 vs. 38-40/≥41 ** 35-37 vs ≥41 *

- Total doses of FSH and days of stimulation increased after age 37 while number of retrieved oocytes and number of mature oocytes (MII) decreased with advanced age
- No statistical differences between antagonist versus progestin-primed ovarian stimulation (PPOS) protocols for each age group

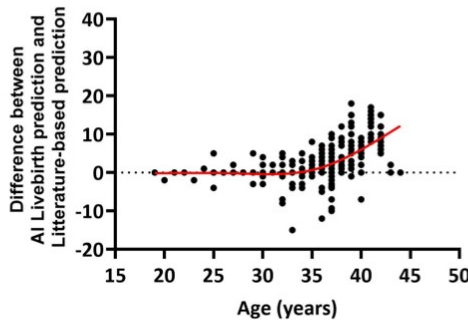
AI predictions



RDI = Relative Dispersion Index calculated as the IQR-to-median ratio

- Median AI **blastulation** scores per patient were **48.0** (IQR: 44.0-50.9)% (<35), **46.9** (43.0-49.9)% (35-37), **44.7** (40.2-49.3)% (38-40), and **43.7** (39.0-47.2)% (≥41) ; **relative dispersion index** (IQR-to-median ratio, reflection of **variability**), of **14.4%**, **14.7%**, **20.4%** and **18.9%**, respectively
- For **AI-based livebirth** prediction scores, medians were **69.5** (55.0-77.0)% (<35), **63.0** (47.7-71.0)% (35-37), **35.5** (21.5-56.2)% (38-40), and **28.5** (19.0-38.5)% (≥41); with respective **variability** coefficients of **31.6%**, **37.0%**, **97.7%** and **68.4%**
- For **literature livebirth** prediction scores, medians were **69.5** (50.5- 80.4)% (<35), **60.0** (43.5-70.5)% (35-37), **26.0** (17.5-46.5)% (38-40), **16.2** (10.2-23.9)% (≥41); with respective **variability** coefficients of **43.0%**, **45.0%**, **111.5%** and **83.9%**

Discrepancies between AI predictions and standard literature based-predictions



- Proportion of patients whose **AI-predicted outcomes deviated from standard literature-based predictions** (which rely solely on age and number of oocytes retrieved, excluding morphological features) **increased with age**

	<35 n=122	35-37 n=126	38-40 n=52	≥41 n=26	P-value
AI LB prediction comparable to literature	79 (64.8%)	64 (50.8%)	4 (7.7%)	4 (15.4%)	35 vs 38-40/≥41 *** 35-37 vs 38-40/≥41 *** & **
AI LB prediction different to literature	43 (35.2%)	62 (49.2%)	48 (92.3%)	22 (84.6%)	
Lower than the literature prediction	22 (51.2%)	49 (79.0%)	2 (4.2%)	22 (100%)	
Higher than the literature prediction	21 (48.8%)	13 (21%)	46 (95.8%)	0 (0%)	

- Median differences** between AI livebirth predictions and literature also **increased with age**
- This findings suggest **relevant contribution of AI-based model in older patients** supporting the hypothesis of **increased variability in oocyte quality with age**

	AI LB prediction (%)	Literature LB prediction (%)	AI LB – Literature LB (%)	P-value
<35, n=122	64.4 ± 18.9 69.5 (55.0, 77.0)	64.6 ± 21.4 69.5 (50.5, 80.4)	-0.12 ± 5.2 -3.6 (-0.5, 3.5)	NS
35-37, n=126	58.6 ± 19.4 63.0 (47.7, 71.0)	56.3 ± 21.0 60.0 (43.5, 70.5)	2.3 ± 5.0 3.0 (-1.0, 6.0)	***
38-40, n=52	38.5 ± 20.2 35.5 (21.5, 56.2)	31.4 ± 19.3 26.0 (17.5, 46.5)	7.1 ± 5.1 7.5 (3.6, 9.9)	***
≥41, n=26	28.3 ± 13.1 28.5 (19.0, 38.5)	18.0 ± 11.3 16.2 (10.2, 23.9)	10.2 ± 6.0 10.7 (5.9, 16.0)	***

Data are expressed as mean ± SD, median (IQR25-75), or n (%)

CONCLUSIONS

The application of AI-based models for the assessment of oocyte quality remains in its early stages. Although promising, their integration into clinical practice requires overcoming several methodological and practical challenges. Our results suggest that oocyte quality becomes increasingly variable with age, highlighting the potential utility of AI tools as personalized counseling aids, particularly for patients considering elective fertility preservation at older ages. The divergence between AI predictions and estimates from the existing literature becomes clinically significant after the age of 37 (with a mean difference exceeding 7%). These findings support that AI-models provide novel and relevant insights beyond those reported by the scientific literature and may be particularly valuable for counseling patients over 37 years of age.

REFERENCES

- Boylan CF et al.** Ex ovo omnia-why don't we know more about egg quality via imaging? *Biol Reprod.* 2024;*110*(6):1201-12.
- Fjeldstad J et al.** An artificial intelligence tool predicts blastocyst development from static images of fresh mature oocytes. *Reprod Biomed Online.* 2024;*48*(6):103842.
- Cimadomo D et al.** Impact of Maternal Age on Oocyte and Embryo Competence. *Front Endocrinol (Lausanne).* 2018;*9*:327.
- Iannone A et al.** On the role of artificial intelligence in analysing oocytes during in vitro fertilisation procedures. *Artif Intell Med.* 2024;*157*:102997.
- Hall JMM et al.** Use of federated learning to develop an artificial intelligence model predicting usable blastocyst formation from pre-ICSI oocyte images. *Reprod Biomed Online.* 2024;*49*(6):104403.

ACKNOWLEDGEMENTS

The authors acknowledge the support and collaboration of Future Fertility for data collection.

The authors have no conflict of interest to declare.

CONTACT INFORMATION

I.PAQUET, MD, Université de Montréal, Montréal, QC, Canada, i.paquet@cliniqueovo.com

L.LAPENSEE, MD, Université de Montréal, Ovo Clinique, Montréal, QC, Canada,

M. CHOW-SHI-YEE, Ovo Clinique, Montréal, QC, Canada

S. PHILLIPS, Université de Montréal, Ovo Clinique, Montréal, QC, Canada