

ENHANCING CLINICAL EFFICIENCY AND GUIDED CARE IN FERTILITY CLINICS THROUGH AI INTEGRATION: A TRAINING APPROACH AND OUTCOME EVALUATION

Authors

Di-Castro Shashua Shirli, Averbuch Dana, Israel Ravid, Heilman Gabrielle

Affiliation

Embie Clinic

Background

Integration of generative artificial intelligence (GenAI) tools in reproductive healthcare has the potential to significantly enhance clinical efficiency and decision-making. However, the adoption of these tools introduces uncertainty, as care teams in fertility clinics may not be accustomed to working with them. Thus, it is crucial to approach AI-generated results with critical thinking, making the training of care teams to assess and utilize these tools essential.

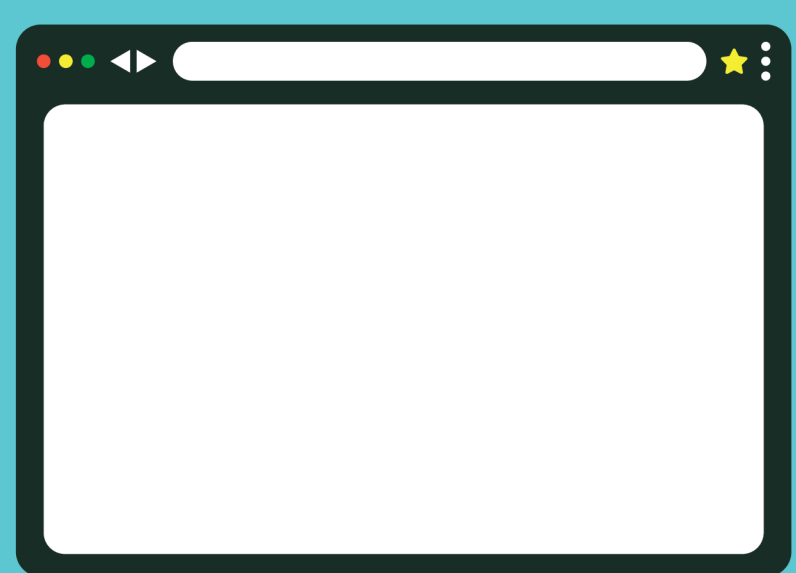
Objective

This study aimed to train care teams to effectively employ a GenAI tool designed to summarize clinical records from fertility clinics into progress notes. We guided the care teams by providing instructions on evaluating AI outputs and analyzing GenAI-generated summaries in comparison to actual clinical data.

Materials and methods

In January 2024, the study evaluated the performance of GenAI in retrospectively generating 85 patient progress note summaries across four categories: activity notes (63), patient-staff communications (3), consultation notes (14), and cycle summaries (5). GPT-4, integrated into Embie, the fertility clinic's electronic health system (EHR) was tasked with summarizing data narratively while incorporating key event details.

The training targeted six care team members at one clinic, consisting of nurses, administrative staff and embryologists, along with two independent domain experts (DE). Trainees compared GenAI-generated summaries with actual EHR data, assessing content reliability and satisfaction, on a scale from 1 to 5. They also provided free-text feedback.



CLINICAL TEAM ENTERS DATA INTO EMR

1

Add follicular scan

DATE OF EXAMINATION

10/11/2023

AFC LEFT

5

AFC RIGHT

6

ENDOMETRIUM

8

LEFT FOLLICLES

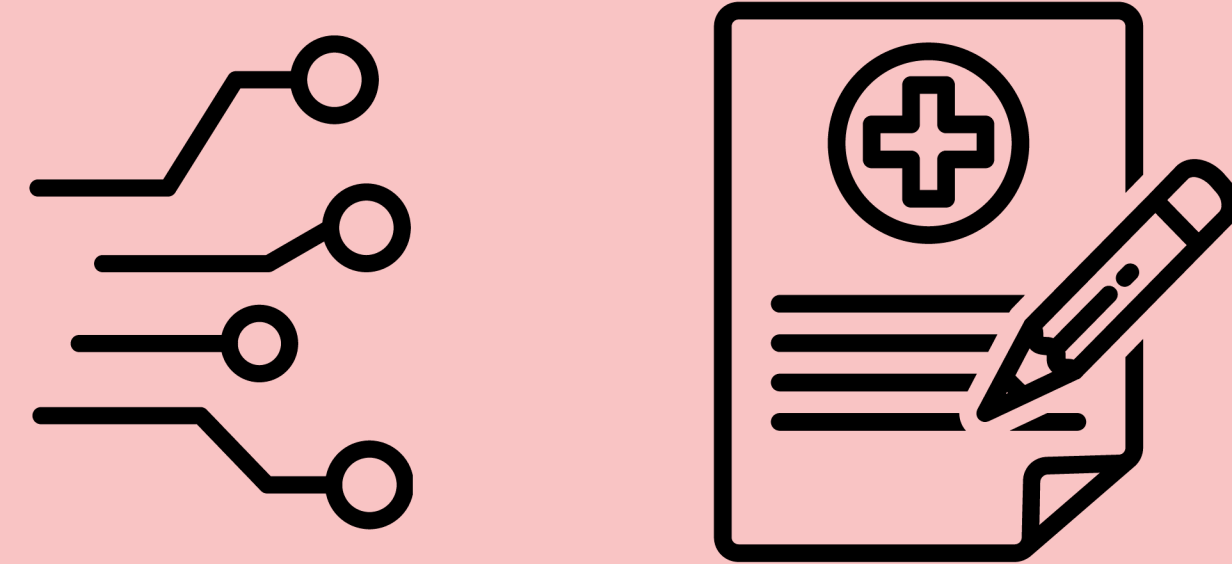
Left Follicles

RIGHT FOLLICLES

Right Follicles

PROCEDURE NOTES

Bloods taken, rescan Fri, continue Menopur 300iu and Fyremadel



LLM GENERATES SUMMARY

2

On October 11, 2023, the patient underwent a follicular scan as part of their in-cycle monitoring. Blood samples were collected during the visit. The patient has been instructed to return for a rescan on Friday and to continue the current medication regimen, which includes Menopur at a dosage of 300 international units and Fyremadel. The scan revealed that the patient has 6 antral follicles on the right ovary and 5 antral follicles on the left ovary, although the sizes of these follicles were not documented in the notes. The endometrial thickness was measured at 8mm.



CLINICAL TEAM RATES SUMMARY OUTPUT

3

Above you will find the Activity feed for patients at your fertility clinic. The Embie Scribe (GenAI) created summaries of activities based on events and actions taken within Embie during the duration of the patient's treatment in the clinic.

Please rate each summary per the instructions below:

1. Indicate how **reliable** the genAI summary is between 1-5, where 1 is not reliable and 5 is very reliable
2. Indicate how **satisfied** you are with the genAI summary between 1-5, where 1- not satisfied and 5 - very satisfied
3. Write your **comments** about the summary. You can write anything you think is important to notice.

Results

A total of 300 ratings were collected regarding content reliability and satisfaction. The average collective score was 4.5, with content reliability and content satisfaction rated at 4.69 and 4.31 respectively. Content reliability scores indicated high reliability from trainees, while content satisfaction scores showed more variability, with domain expert scoring 4.43, and care teams expressing a more cautious perspective at 3.83.

Qualitative feedback revealed concerns regarding GenAI's tendency to include unnecessary opinions, redundant details, and excessive verbosity. Furthermore, training enabled participants to identify key discrepancies and inaccuracies in data entries, attributing some issues to human error rather than GenAI's limitations. This highlights GenAI potential as a quality control measure and emphasizes the importance of accurate data input for optimal AI performance.

Average Score

4.5



4.31

Satisfaction Score

4.69

Reliability Score

Conclusion

The findings emphasize the necessity of tailored training programs for care teams to effectively collaborate with GenAI tools in clinical environments. By empowering care teams with the necessary skills to evaluate and critique AI-generated outputs, we can enhance the potential for GenAI tools to improve patient care outcomes in fertility clinics and beyond.