Received: 13 January 2024 | Accepted: 16 August 2024

DOI: 10.1002/rmb2.12603

Reproductive Medicine and Biology

WILEY

ORIGINAL ARTICLE

Development of an AI-based support system for controlled ovarian stimulation

Yoshimasa Asada[1,2] | Tomoya Shinohara[3] | Sho Yonezawa[3] | Tomoki Kinugawa[1] | Emiko Asano[1,2] | Masae Kojima[1,2] | Noritaka Fukunaga[1,2] | Natsuka Hashizume[3] | Yoshiki Hashiba[1] | Daichi Inoue[1] | Rie Mizuno[1] | Masaya Saito[1] | Yoshinori Kabeya[3]

[1]Asada Ladies Clinic, Nagoya, Japan

[2]Asada Institute for Reproductive Medicine, Kasugai, Japan

[3]IBM Japan Ltd., Tokyo, Japan

Correspondence
Yoshimasa Asada, Asada Ladies Clinic, Nagoya, Japan.
Email: y_asada@ivf-asada.jp

Abstract

Purpose: Controlled ovarian stimulation (COS) is vital for IVF. We have developed an AI system to support the implementation of COS protocols in our clinical group.

Methods: We developed two models as AI algorithms of the AI system. One was the oocyte retrieval decision model, to determine the timing of oocyte retrieval, and the other was the prescription inference model, to provide a prescription similar to that of an expert physician. Data was obtained from IVF treatment records from the In Vitro Fertilization (IVF) management system at the Asada Ladies Clinic, and these models were trained with this data.

Results: The oocyte retrieval decision model achieved superior sensitivity and specificity with 0.964 area under the curve (AUC). The prescription inference model achieved an AUC value of 0.948. Four models, namely the hCG prediction model, the hMG prediction model, the Cetrorelix prediction model, and the Estradiol prediction model included in the prescription inference model, achieved AUC values of 0.914, 0.937, 0.966, and 0.976, respectively.

Conclusion: The AI algorithm achieved high accuracy and was confirmed to be useful. The AI system has now been implemented as a COS tool in our clinical group for self-funded treatments.

KEYWORDS
artificial intelligence, in vitro fertilization, oocyte retrieval, ovulation induction, prescriptions

1 | INTRODUCTION

Successful IVF requires multiple mature oocytes. The number of oocytes required for a single live birth from a 2012 paper was reported to average 25.1, considering maturity rate, normal fertilization rate, blastocyst arrival rate, implantation rate, and miscarriage rate.[1] More recently Polyzos et al it has been clearly shown that, the retrieval of multiple oocytes from one ovarian stimulation cycle can increase the cumulative live birth rate after repeated cycles and shorten the time to conception.[2,3] It was also reported that the pregnancy rate of frozen oocytes was higher than that of oocytes that were returned directly to the mother.[4] The NICE (National Institute for Health and Care Excellence) guidelines favor COS over spontaneous cycles.[5] Therefore, COS is a crucial step in any IVF/ICSI treatment.

However, in order to manage COS properly a high degree of training and experience of the physician are required. The clinical team at the Asada Ladies Clinic, have been accumulating COS methods for over 30 years. Transferring these skills to younger physicians has become a challenge in the clinic. We have observed a discrepancy in the success rates between experienced physicians and younger physicians as well as experienced physicians' spending a great deal of time in monitoring the techniques of younger physicians.

The recent development of artificial intelligence (AI) technology is remarkable its use in the fertility treatment field is increasing.[6–8] For example, Khosravi developed STORK, an AI tool that evaluated embryo quality from human embryo images.[5] Goyal developed an AI tool that predicted live birth based on patient background information.[6]

Here we describe the use of AI for the development of a tool to manage COS that would assist physicians in real-time and to meet the needs of younger physicians' education and reduce the managerial workload of expert physicians at the Asada Ladies Clinic.

2 | MATERIALS AND METHODS

2.1 | Data

The data for training and evaluation of AI tool were the IVF records in Asada Ladies Clinic's IVF management system. The IVF management system contains the examination records and procedures of the patients including the number and size of follicles and mature follicles identified by ovarian ultrasound, hormone levels, follicle stimulating hormone (FSH), luteinizing hormone (LH), estradiol (E_2), human chorionic gonadotropin (hCG), progesterone (P_4), etc. These clinical laboratory data, excluding AMH, are collected during each patient visit four to six times at intervals of 2–5 days, starting before the onset of ovarian stimulation. AMH is measured before the initial treatment and is retested 6 months to 1 year later.

2.2 | Architecture

The AI-Based Support System for Controlled Ovarian Stimulation, named AACS (the AI Asada-style controlled ovarian stimulation support system), consisted of two AI models (Figure 1A). One is an oocyte retrieval decision model, and the other is a prescription inference model. The oocyte retrieval decision model infers the timing of oocyte retrieval based on the patient's follicles and hormone status in the same way as an expert physician. The prescription inference model is a model that provides a drug prescription similar to that of an expert physician based on the patient's follicles and hormone status.

The process was as follows. After the patient's information was inputted, it was evaluated in the oocyte retrieval decision model, which determined whether or not to retrieve oocytes. If the optimal timing for retrieving mature oocytes is achieved, the process is completed with a recommendation to the physician for oocyte retrieval. If not, the patient data is passed to the prescription inference model, and the prescription inferred by the model is presented to the physician (Figure 1A).

The prescription inference model consisted of four components: the hMG prediction model, the hCG prediction model, the Cetrorelix prediction model, and the Estradiol prediction model. The hMG prediction model infers which a human menopausal gonadotropin (hMG) preparation (HMG for injection [FERRING]/Menopur®; Ferring Pharmaceuticals, Tokyo, Japan) or a recombinant follicle-stimulating hormone (FSH) preparation (Gonalef®; Merck Biopharma, Tokyo, Japan) should be prescribed and the dosage. The hCG prediction model infers whether or not to prescribe hCG preparation (GONATOROPIN®; ASKA Pharmaceutical, Tokyo, Japan) in addition to HMG for injection [FERRING] in amounts of either 30 or 50 units to enhance LH activity and the dosage. The Cetrorelix prediction model infers whether or not to prescribe a GnRH antagonist preparation (Cetrotide®; Merck Biopharma, Tokyo, Japan). The Estradiol prediction model infers whether or not to prescribe oral estradiol preparation (Julina®, Bayer Yakuhin, Osaka, Japan) to increase E_2 levels at the start of stimulation (Figure 1B).

2.3 | Training and evaluating the oocyte retrieval decision model

IVF data recorded from June 2017 to November 2021 (before the changes to Japan's health insurance coverage in April 2022) were used to train and evaluate the oocyte retrieval decision model. The total number of patients was 5969, and the number of cycles was 7850. We extracted records performed by an expert physician from the IVF data. The number of patients and cycles and examinations performed by the expert physician was 971, 1068 and 1345, respectively. Of these, 316 examinations whose next examination date was the oocyte retrieval date were labeled “Oocyte retrieval”, and 1029 examinations whose next examination date was not the oocyte retrieval date were labeled “Continued stimulation”. The output value of the model ranges from 0.0 to 1.0. We defined it as oocyte retrieval if the value was 0.5 or more and as continued stimulation if the value was less than 0.5.

The input information for the oocyte retrieval decision model was the total number of follicles in the patient's right and left ovaries, the average of the maximum diameter and the diameter perpendicular to it for both mature and immature follicles observed (up to 4) in the right and left ovary, FSH, LH, E_2, hCG, P_4, age, Anti-Müllerian hormone: AMH, the number of days since the start of stimulation, the expected number of mature follicles, and size of mature follicles. As additional features, we used the difference between the number of immature and mature follicles and the mean diameter squared follicle size. We trained and evaluated the oocyte retrieval decision model with the features using the 5-fold cross-validation.[9]

A cross-validation is an established method used to statistically evaluate models and assess how accurately the models predict

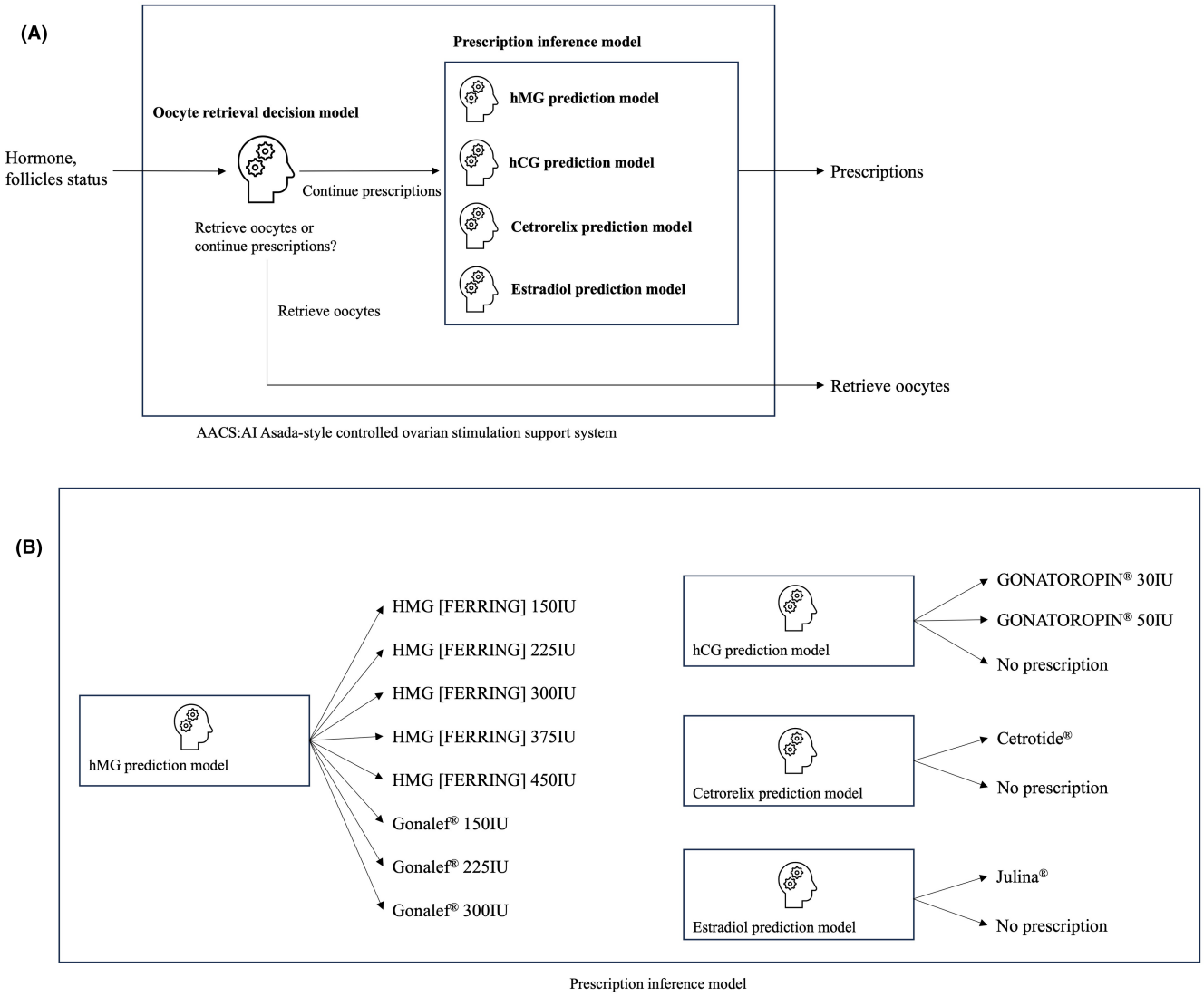

FIGURE 1 AACS architecture. (A) AACS consists of the oocyte retrieval decision model and the prescription inference model. (B) The prescription model has four models: The hMG prediction model, the hCG prediction model, the Cetrorelix prediction model, and the Estradiol prediction model.

unknown data. A cross-validation randomly divides the objective data into training and test sets, which are used for training and evaluation, respectively. A 5-fold cross-validation divides the samples into five groups. One group is used for a test set, and four groups are used for a training set. The test can be performed five times by shifting the test set each time, and the results of all test lots are combined and averaged. In this way, all the data is effectively used for both training and testing, and the evaluation results can be statistically stable and accurate.

2.4 | Training and evaluating the prescription inference model

The training and evaluation data for the prescription inference model were generated by re-reviewing IVF data recorded between June 2017 and November 2021 with an expert physician. The reason for the re-review was that the approach to prescribing is constantly evolving, and it was necessary to align with the latest practices of the expert physicians. As a result, there were 374 patients, 434 cycles, and 1397 exams that were acceptable for inclusion by the expert physicians. This data was used to train four models included in the prescription inference model: the hMG prediction model, the hCG prediction model, the Cetrorelix prediction model, and the Estradiol prediction model. Some product data was not used (GANIREST®, Organon, Tokyo, Japan) (Human Menopausal Gonadotrophin for injection, Fuji Pharma, Toyama, Japan) (HMG Intramuscular Injection, ASKA Pharmaceutical, Tokyo, Japan) because the number of prescriptions was low (less than 10 prescriptions). As a result, the hMG prediction model was designed to predict HMG for injection [FERRING] 150 IU, HMG for injection [FERRING] 225 IU, HMG for injection [FERRING] 300 IU, HMG for injection [FERRING] 375 IU, HMG for injection [FERRING] 450 IU, Gonalef® 150 IU, Gonalef® 225 IU or

Gonalef® 300IU. The hCG prediction model was designed to predict either non-prescription, GONATOROPIN® 30IU or GONATOROPIN® 50IU. The Cetrorelix prediction model predicted whether Cetrotide® should be prescribed or not. The Estradiol prediction model predicted whether Julina® should be prescribed or not (Figure 1B).

The input data for the four models were the number of follicles on each side, the size of follicles on each side (mean), FSH, E_2, hCG, P_4, AMH, age, the number of days since stimulation began, and the drug prescribed at the last diagnosis before oocyte retrieval. In addition, the rate of change of each hormone value since the previous visit, the moving average of each hormone value of 3 tests, the mean mature follicle size, and the variability of mature follicle size were used as additional features. Each of the four models was trained and evaluated using 5-fold cross-validation.

2.5 | Model

We used LightGBM[10] as the machine learning model for both the oocyte retrieval decision model and the prescription inference model. LightGBM employs a gradient boosting framework, which constructs decision trees sequentially, with each new tree helping correct errors made by previously trained trees. LightGBM is also recognized for its exceptionally high prediction accuracy and computational efficiency and is used in analyses involving medical data.[11] This study opted for LightGBM due to its superior performance over alternative methodologies, such as support vector machine,[12] logistic regression, and deep neural networks in the pre-validation phase. In the pre-validation phase, the accuracy of LightGBM, support vector machine, logistic regression, and deep neural networks were recorded at 0.951, 0.936, 0.936, and 0.932, respectively. We used a Python library, LightGBM, version 4.0.0.5, for developing the AI model.

2.6 | Metrics

The study used two metrics to evaluate the AI algorithms. One was accuracy, and the other was an area under the curve (AUC) by the ROC (Receiver Operating Characteristic) curve.

1. Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

2. ROC and AUC

The ROC curve is a graphical curve that represents the true positive fraction (=TPF) and false positive fraction (=FPF) calculated and plotted on a plane with TPF on the vertical axis and FPF on the horizontal axis and connected by a line.

When the ROC curve is created, the area under the graph is called the Area Under the Curve (AUC). The AUC takes values from 0 to 1, with values closer to 1 indicating a higher discrimination capacity.

If the model has three or more outputs, multiple ROC curves are created by considering them as a binary classification of the target and the others. In such cases, the study used an average AUC calculated to average all AUC as an evaluation index.

3. Feature Importance in LightGBM[10]

Feature importance in LightGBM is a metric that indicates the relative significance of each feature in making predictions. It is determined through the model's training process, where the algorithm evaluates the contribution of each feature to the model's accuracy. Feature importance is calculated by considering the extent to which each feature contributes to the reduction of the loss function across all trees within the model. Understanding feature importance helps with model interpretability, allowing us to see which factors drive the model's decisions.

3 | RESULTS

The AUC and accuracy of the oocyte retrieval decision model were 0.964% and 96.9%, respectively (Figure 2A). The confusion matrix also indicated high accuracy for this model (Figure 2B). Upon examining the feature importance of the predictive model, we found that the second largest follicle size in the right ovary (2nd_Lg_Foll_size_R) and the maximum diameter of the largest follicle in the left ovary (Lg_Foll_Diam_L) had importance scores of 0.53 and 0.14, respectively (Figure 2C). The 2nd_Lg_Foll_size_R refers to the size calculated using the maximum diameter and its perpendicular diameter for the second largest follicle in the right ovary during the IVF, while the Lg_Foll_Diam_L denotes the maximum follicle diameter recorded in the first field for the left ovary. These results suggest that follicle size, particularly the diameter, is a crucial factor in the decision-making process for oocyte retrieval.

The average AUC and the accuracy of the prescription inference model were 0.948% and 86.9%. The four models; the hMG prediction model, the hCG prediction model, the Cetrorelix prediction model, and Estradiol prediction model, which were included in the prescription inference model, achieved an average AUC 0.914 (accuracy 71.9%), an average AUC 0.937 (accuracy 93.5%), an AUC 0.966 (accuracy 88.6%) and an AUC 0.976 (accuracy 93.6%), respectively (Figures 3A, 4A, 5A, 6A). Figures 3B, 4B, 5B and 6B show the confusion matrices of those results.

The accuracy at which all four model outputs were correct was 56.2%, and the accuracy was 76.2% when we regarded that the predicted drugs were the same as the actual drugs as the correct answers even if the amount was different.

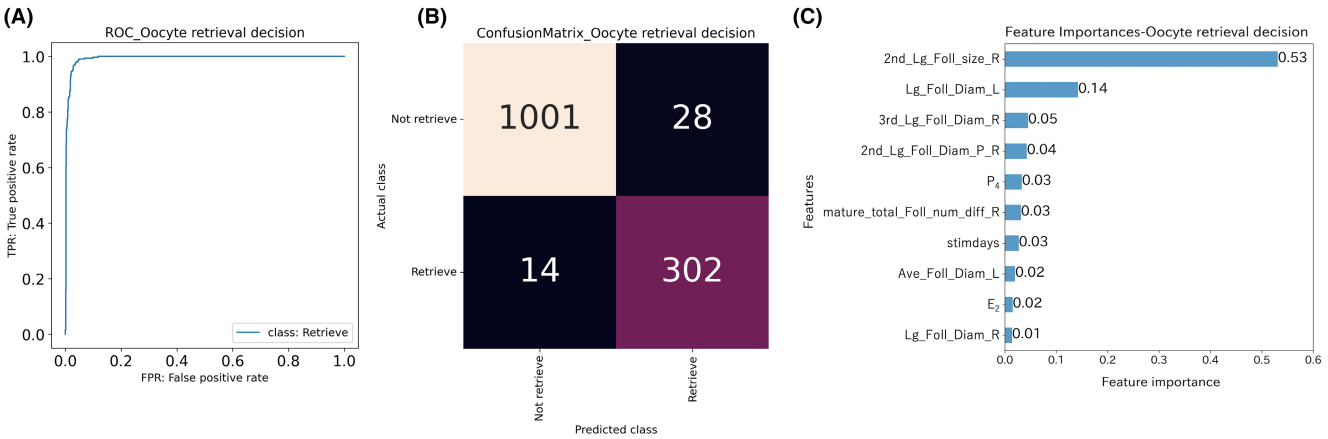

FIGURE 2 Evaluation of the oocyte retrieval decision model. (A) is the ROC curve of the oocyte retrieval decision model. (B) is the confusion matrix of the model. A confusion matrix is used to evaluate the performance of classification models. (C) is the feature importance of the oocyte retrieval decision model. The 2nd_Lg_Foll_size_R refers to the size calculated using the maximum diameter and its perpendicular diameter for the second largest follicle in the right ovary during the IVF procedure. The Lg_Foll_Diam_L represents the maximum follicle diameter for the largest follicle in the left ovary. The 3rd_Lg_Foll_Diam_R denotes the maximum follicle diameter for the third largest follicle in the right ovary. The 2nd_Lg_Foll_Diam_P_R refers to the diameter perpendicular to the maximum diameter of the second largest follicle in the right ovary. P_4 and E_2 represent the blood levels of progesterone and estrogen, respectively. The mature_total_Foll_num_diff_R is the calculated difference between the number of mature and immature follicles in the right ovary. The stimdays is the number of days since stimulation began. The Ave_Foll_Diam_L represents the average of the maximum diameter and its perpendicular diameter for the follicles in the left ovary. The Lg_Foll_Diam_R refers to the maximum follicle diameter for the largest follicle in the right ovary.

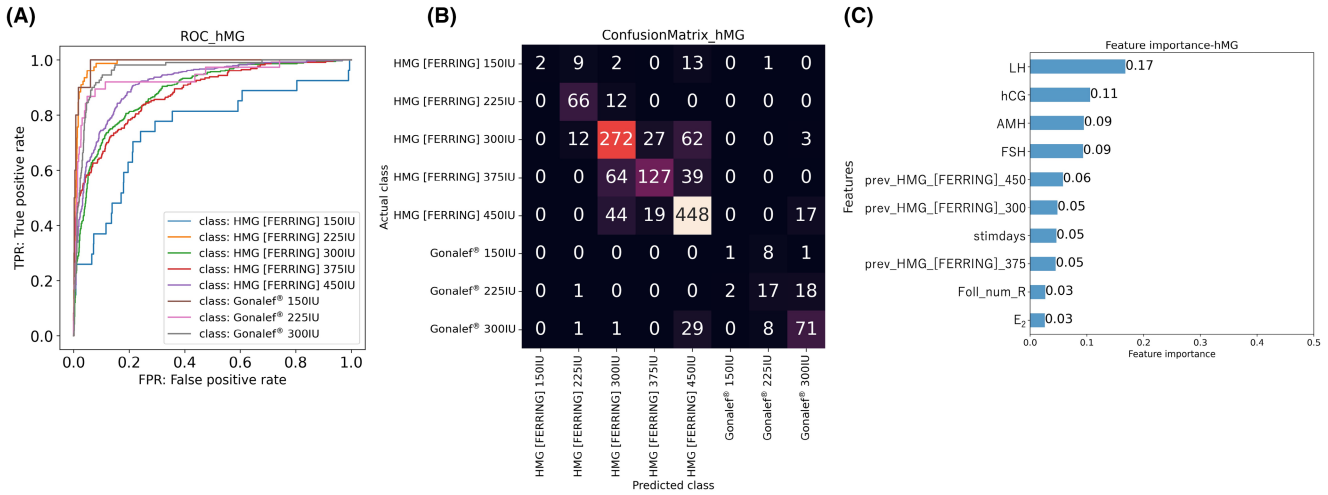

FIGURE 3 Evaluation of the hMG prediction model. (A) is the ROC curve of the hMG prediction model. (B) is the confusion matrix of the hMG prediction model. (C) is the feature importance of the hMG prediction model. The LH, hCG, AMH, FSH features represent the levels of hormones measured during the IVF procedure. The features starting with "prev_" represent the medications prescribed during the previous examination. The Foll_num_R represents the number of follicles observed in the right ovary.

The AUCs of HMG for injection [FERRING] 150IU, HMG for injection [FERRING] 225IU, HMG for injection [FERRING] 300IU, HMG for injection [FERRING] 375IU, HMG for injection[FERRING] 450IU, Gonalef® 150IU, Gonalef® 225IU and Gonalef® 300IU as the hMG prediction model outputs were 0.753, 0.987, 0.887, 0.881, 0.915, 0.989, 0.940, and 0.963, respectively.

The AUC of GONATOROPIN® 30IU, GONATOROPIN® 50IU, and non-prescription as the hCG prediction model outputs were 0.964, 0.867, and 0.981, respectively. It was difficult for the model to answer GONATOROPIN® 50IU. One potential explanation could be the low rate of prescriptions (Figure 4A).

Upon evaluating the feature importance for each model, it became apparent that hormone levels at the time of examination, such as LH and hCG, are crucial across all four models. LH ranks as the most important feature for both the hMG and hCG prediction models, and it is the second most important feature for the Cetrorelix and Estradiol

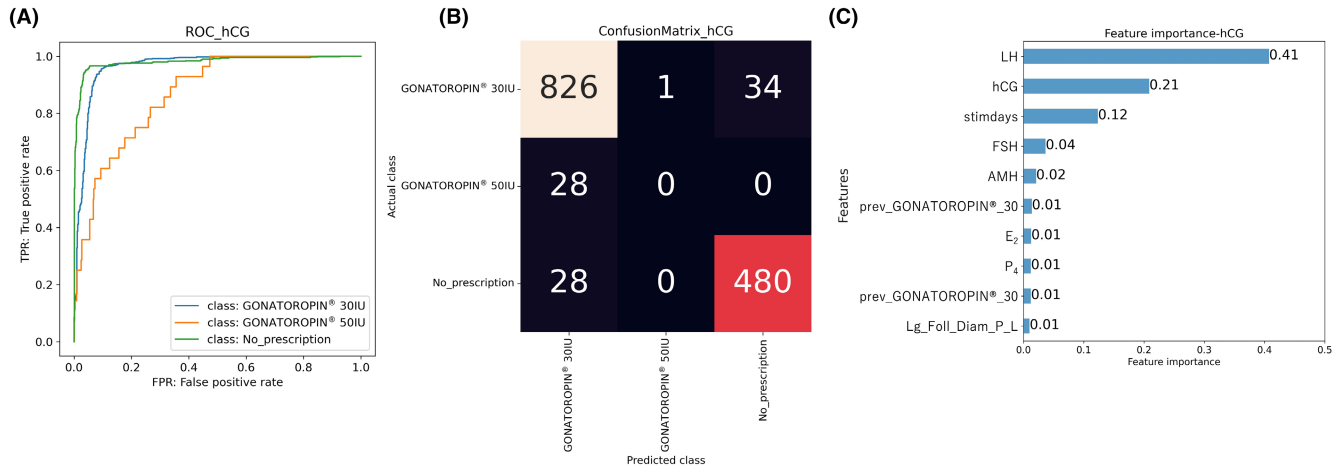

FIGURE 4 Evaluation of the hCG prediction model. (A) is the ROC curve of the hCG prediction model. (B) is the confusion matrix of the hCG prediction model. (C) is the feature importance of the hCG prediction model. The LH, hCG, FSH, AMH, E_2, P_4, features represent the levels of hormones measured during the IVF procedure. The stimdays is the number of days since stimulation began. The features starting with "prev_" represent the medications prescribed during the previous examination. The Lg_Foll_Diam_P_L represents the diameter perpendicular to the maximum diameter of the largest follicle in the left ovary.

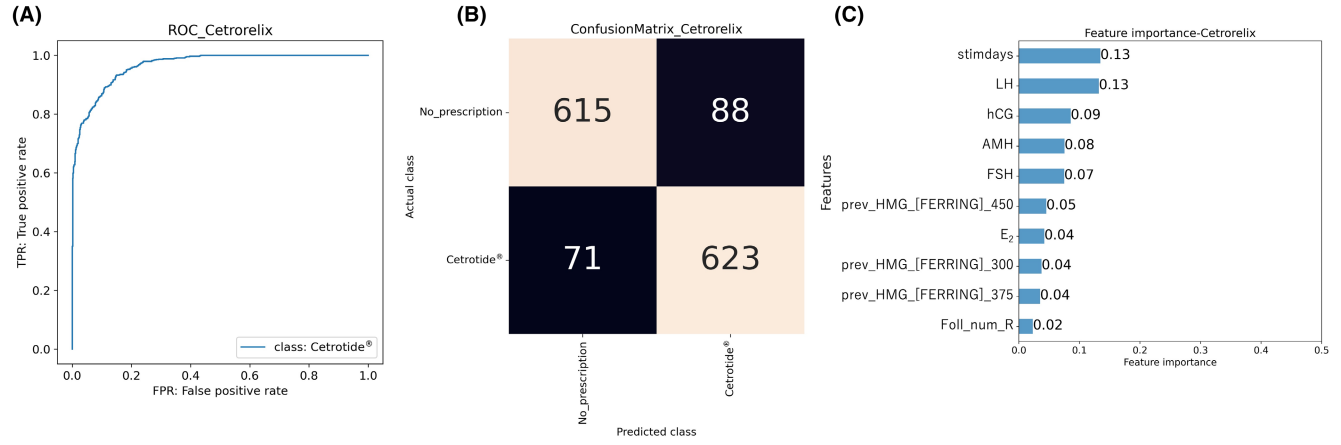

FIGURE 5 Evaluation of the Cetrorelix prediction model. (A) is the ROC curve of the Cetrorelix prediction model. (B) is the confusion matrix of the Cetrorelix prediction model. (C) is the feature importance of the Cetrorelix prediction model. The stimdays is the number of days since stimulation began. The LH, hCG, AMH, FSH, E_2 features represent the levels of hormones measured during the IVF procedure. The features starting with "prev_" represent the medications prescribed during the previous examination. The Foll_num_R represents the number of follicles observed in the right ovary.

models (Figures 3C, 4C, 5C, 6C). Similarly, hCG ranks as the second most important feature for the hMG and hCG prediction models, while it is the third and fourth most important feature for the Cetrorelix and Estradiol models, respectively (Figures 3C, 4C, 5C, 6C).

Notably, in the fundamental hMG prediction model, AMH with a feature importance of 0.09 ranks third (Figure 3C), highlighting its significance. In the other three prescription models, the number of days since the start of stimulation (stimdays) ranked highly, underscoring the importance of the timing within the stimulation cycle for decision-making. This is especially true for the Cetrorelix and Estradiol prediction models, where the stimulation cycle is the most critical factor, with the number of days since the start of stimulation (stimdays) ranking as the most important feature in both models (Figures 5C and 6C).

These findings emphasize the pivotal role of hormonal levels, particularly LH and hCG, in guiding the prescription decision-making process across all four models. Furthermore, the prominence of AMH in the hMG model and the high ranking of the number of days since the start of stimulation (stimdays) in the Cetrorelix and Estradiol models provide valuable insights into the specific factors that influence prescription decisions at different stages of the assisted reproductive technology process.

4 | DISCUSSION

In this study, we developed AACS, an AI to guide and support ovarian stimulation in our clinical situation. AACS consists of an oocyte

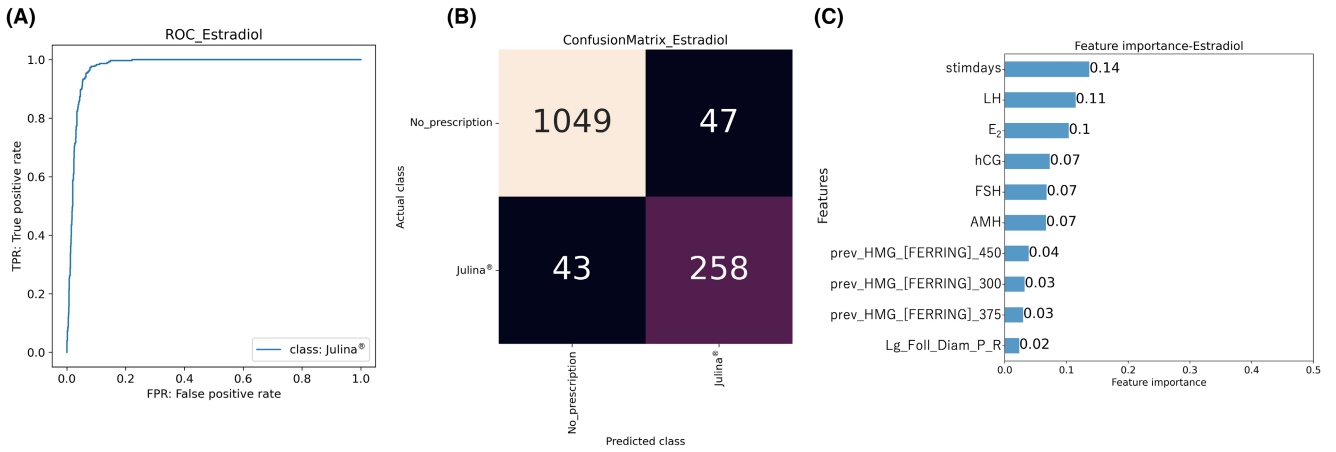

FIGURE 6 Evaluation of the Estradiol prediction model. (A) is the ROC curve of the Estradiol prediction model. (B) is the confusion matrix of the Estradiol prediction model. (C) is the feature importance of the Estradiol prediction model. The stimdays is the number of days since stimulation began. The LH, E_2, hCG, FSH, AMH features represent the levels of hormones measured during the IVF procedure. The features starting with "prev_" represent the medications prescribed during the previous examination. The Lg_Foll_Diam_P_R represents the diameter perpendicular to the maximum diameter of the largest follicle in the right ovary.

retrieval decision model to help determine when to retrieve oocytes and a prescription inference model to present appropriate prescriptions based on patient conditions. The results of this study confirmed the usefulness of the AACS, with both models achieving high accuracy.

In addition, our findings emphasize the importance of follicle diameter in the oocyte retrieval decision-making process and LH and hCG in the prescription decision-making process. LH is crucial for preventing premature luteinization and determining the duration of cetrorelix administration to suppress the LH surge. The decrease in LH associated with an increase in E_2 can adversely affect oocyte maturation; therefore, a low dose of usually 30 IU of hCG is administered not as a typical trigger but to supplement LH activity to prevent its decline. This is followed by monitoring the effects on oocyte maturation based on LH and hCG blood levels.

The AACS has now been integrated into the IVF management system at the Asada Ladies Clinic to train younger physicians. However, due to Japan's health insurance coverage changes for infertility treatment implemented in April 2022, insurance was suddenly expanded to cover assisted reproductive treatments, imposing significant restrictions on treatment, such as limitations on the drugs used and their dosage and administration, as well as restrictions on the number of tests. This has led to a reduction in patients receiving treatment, however we exepct that this will improve in the future.

A future challenge is the prediction of the date for oocyte retrieval. The oocyte retrieval decision model in AACS is designed to infer the oocyte retrieval date by a few days before the oocyte retrieval date, but not earlier than that; being able to infer the oocyte retrieval date 2 to 3 weeks in advance may be a future requirement for some patients, who wish to balance work and infertility treatment.

AACS is only compatible with ovarian stimulation using the antagonist, and future research is needed to make it compatible with other treatment methods, such as agonists and simple mild stimulation protocols.

In addition, in order to facilitate inquiry response both from patients as well as trainee clinicians the usefulness of the large language model, which is another AI technology and utilized in the chat GPT[13] is expected.

ACKNOWLEDGMENTS

We would like to express our deepest gratitude to the members of our institution for their diligent and thorough work in organizing and preparing the extensive data that was crucial for this study. Additionally, we extend our sincere thanks to the members at IBM, our co-authors, for their precise and skilled handling of data processing. Their expertise and collaborative efforts greatly contributed to the quality and accuracy of our findings.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

HUMAN RIGHTS STATEMENTS AND INFORMED CONSENT

All the procedures were followed in accordance with the ethical standards of the responsible committees on human experimentation (institutional and national) and with the principles of the Helsinki Declaration of 1964 and its later amendments. This is a retrospective study in patients who submitted informed consent for undergoing fertility treatment at our IVF center.

ANIMAL STUDIES

This article does not contain any study with animal participants that have been performed by any of the authors.

ORCID

Yoshimasa Asada https://orcid.org/0000-0002-0448-9953
Noritaka Fukunaga https://orcid.org/0000-0003-2161-3319
Yoshinori Kabeya https://orcid.org/0000-0001-5843-4889

REFERENCES

1. Gianaroli L, Racowsky C, Geraedts J, Cedars M, Makrigiannakis A, Lobo R. Best practices of ASRM and ESHRE: a journey through reproductive medicine. Hum Reprod. 2012;27(12):3365–79.
2. Polyzos NP, Drakopoulos P, Parra J, Pellicer A, Santos-Ribeiro S, Tournaye H, et al. Cumulative live birth rates according to the number of oocytes retrieved after the first ovarian stimulation for in vitro fertilization/intracytoplasmic sperm injection: a multicenter multinational analysis including ~15,000 women. Fertil Steril. 2018;110(4):661–670.e1. https://doi.org/10.1016/j.fertnstert.2018.04.039
3. Vaughan DA, Leung A, Resetkova N, Ruthazer R, Penzias AS, Sakkas D, et al. How many oocytes are optimal to achieve multiple live births with one stimulation cycle? The one-and-done approach. Fertil Steril. 2017;107(2):397–404.e3.
4. Asada Y, Tokoro M, Sonohara M, Fukunaga N, Hattori Y, Hashiba Y. Long-term outcomes of freeze-all strategy: a retrospective analysis from a single ART center in Japan. Reprod Med Biol. 2019;18(2):173–9.
5. National Institute for Health and Care Excellence. Fertility problems: assessment and treatment. Published: 2013 Feb. https://www.nice.org.uk/guidance/cg156/chapter/recommendations
6. Khosravi P, Kazemi E, Zhan Q, Malmsten JE, Toschi M, Zisimopoulos P, et al. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. NPJ Digit Med. 2019;2:21.
7. Goyal A, Kuchana M, Ayyagari KPR. Machine learning predicts live-birth occurrence before in-vitro fertilization treatment. Sci Rep. 2020;10:20925.
8. Fernandez EI, Ferreira AS, Cecílio MHM, Chéles DS, Rebeca Souza CM, Nogueira MFG, et al. Artificial intelligence in the IVF laboratory: overview through the application of different types of algorithms for the classification of reproductive data. J Assist Reprod Genet. 2020;37:2359–76.
9. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. International Joint Conference on Artificial Intelligence 1137–1143 (IJCAI 1995).
10. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: a highly efficient gradient boosting decision tree. 31st Conference on Neural Information Processing Systems (NIPS 2017).
11. Seto H, Oyama A, Kitora S, Toki H, Yamamoto R, Kotoku J, et al. Gradient boosting decision tree becomes more reliable than logistic regression in predicting probability for diabetes with big data. Sci Rep. 2022;12:15889.
12. Boser BE, Boser GIM, Vapnik VN. A training algorithm for optimal margin classifiers. Proceedings of the Fifth Annual Workshop on Computational Learning Theory 1992: 144–52
13. Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal P, et al. Language models are few-shot learners. 33rd Conference on Neural Information Processing Systems (NIPS 2020).

How to cite this article: Asada Y, Shinohara T, Yonezawa S, Kinugawa T, Asano E, Kojima M, et al. Development of an AI-based support system for controlled ovarian stimulation. Reprod Med Biol. 2024;23:e12603. https://doi.org/10.1002/rmb2.12603