ISSN 2644-3260 (online)

Artificial Intelligence in *Medical Imaging*

Artif Intell Med Imaging 2020 September 28; 1(3): 87-107

Published by Baishideng Publishing Group Inc

AIMI

Artificial Intelligence in Medical Imaging

Contents

Bimonthly Volume 1 Number 3 September 28, 2020

Contents

ABOUT COVER

Editorial board member of *Artificial Intelligence in Medical Imaging*, Ying-Kun Guo, MD is a Professor and Deputy Director of the Department of Radiology, West China Second University Hospital, Sichuan University (China). His career research has focused on cardiovascular imaging, pediatric radiology, and molecular and metabolic imaging. As an experienced radiologist, Dr. Guo's research achievements have been published in international professional journals, such as *JACC, Radiology, International Journal of Cardiology, Neuroimage, and Journal of Cardiac Magnetic Resonance*. Also, Professor Guo serves as PI of the Molecular Imaging Lab, which was established in 2015 and aims to develop novel molecular imaging in vivo, targeted nano-molecular probe, radiogenomics, and imaging contrast mechanisms, to evaluate emerging MRI sequences in animal models of human diseases, and to translate and apply innovative diagnostic imaging in clinic. (L-Editor: Filipodia)

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Medical Imaging* (*AIMI*, *Artif Intell Med Imaging*) is to provide scholars and readers from various fields of artificial intelligence in medical imaging with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIMI mainly publishes articles reporting research results obtained in the field of artificial intelligence in medical imaging and covering a wide range of topics, including artificial intelligence in radiology, pathology image analysis, endoscopy, molecular imaging, and ultrasonography.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Jia-Hui Li*; **Production Department Director:** *Yun-Xiaojian Wu*; **Editorial Office Director:** *Jin-Lei Wang*.

NAME OF JOURNAL
Artificial Intelligence in Medical Imaging

ISSN
ISSN 2644-3260 (online)

LAUNCH DATE
June 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Xue-Li Chen, Caroline Chung, Ahmed Abdel Khalek Abdel Razek, Jun Shen

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2644-3260/editorialboard.htm

PUBLICATION DATE
September 28, 2020

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

Artificial Intelligence in Medical Imaging

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Med Imaging 2020 September 28; 1(3): 87-93

DOI: 10.35711/aimi.v1.i3.87

ISSN 2644-3260 (online)

EDITORIAL

Current trends of artificial intelligence in cancer imaging

Francesco Verde, Valeria Romeo, Arnaldo Stanzione, Simone Maurea

ORCID number: Francesco Verde 0000-0002-9823-4678; Valeria Romeo 0000-0002-1603-6396; Arnaldo Stanzione 0000-0002-7905-5789; Simone Maurea 0000-0002-8269-3765.

Author contributions: Verde F drafted the manuscript; Romeo V conceptualized and drafted the manuscript; Stanzione A and Maurea S performed critical revision and approved the final manuscript.

Conflict-of-interest statement: The authors declare no conflicts of interest.

Manuscript source: Invited manuscript

Received: August 23, 2020

Peer-review started: August 23,

Francesco Verde, Valeria Romeo, Arnaldo Stanzione, Simone Maurea, Department of Advanced Biomedical Sciences, University of Naples "Federico II", Napoli 80131, Italy

Corresponding author: Valeria Romeo, MD, PhD, Academic Research, Doctor, Research Fellow, Department of Advanced Biomedical Sciences, University of Naples "Federico II", Via S. Pansini 5, Napoli 80131, Italy. valeria.romeo@unina.it

Abstract

In this editorial, we discussed the current research status of artificial intelligence (AI) in Oncology, reviewing the basics of machine learning (ML) and deep learning (DL) techniques and their emerging applications on clinical and imaging cancer workflow. The growing amounts of available "big data" coupled to the increasing computational power have enabled the development of computer-based systems capable to perform advanced tasks in many areas of clinical care, especially in medical imaging. ML is a branch of data science that allows the creation of computer algorithms that can learn and make predictions without prior instructions. DL is a subgroup of artificial neural network algorithms configurated to automatically extract features and perform high-level tasks; convolutional neural networks are the most common DL models used in medical image analysis. AI methods have been proposed in many areas of oncology granting promising results in radiology-based clinical applications. In detail, we explored the emerging applications of AI in oncological risk assessment, lesion detection, characterization, staging, and therapy response. Critical issues such as the lack of reproducibility and generalizability need to be addressed to fully implement AI systems in clinical practice. Nevertheless, AI impact on cancer imaging has been driving the shift of oncology towards a precision diagnostics and personalized cancer treatment.

Key Words: Artificial intelligence; Machine learning; Deep learning; Oncology; Medical imaging; Cancer imaging

Core Tip: Advanced computational systems and availability of multi-dimensional data have led the possibility of artificial intelligence (AI) consisting of machine learning (ML) and deep learning (DL) algorithms to be implemented in healthcare data analysis, with reliable results in the oncology field and particularly in diagnostic imaging tasks. Supervised algorithms are the most common ML models used in medical image analysis, while

2020
First decision: September 13, 2020
Revised: September 22, 2020
Accepted: September 23, 2020
Article in press: September 23, 2020
Published online: September 28, 2020

P-Reviewer: Abdel Razek AAK, Ogino S, Ren J, Wang RF
S-Editor: Wang JL
L-Editor: A
P-Editor: Li JH

convolutional neural networks are the main DL approach. AI-based models have demonstrated outperforming results in oncological risk assessment, lesion detection, segmentation, characterization, staging, and therapy response. Growing emerging evidence supports the leading role of AI in all cancer imaging pathways from screening programs to diagnostic and prognostic tasks, boosting the paradigm of precision medicine.

Citation: Verde F, Romeo V, Stanzione A, Maurea S. Current trends of artificial intelligence in cancer imaging. *Artif Intell Med Imaging* 2020; 1(3): 87-93
URL: https://www.wjgnet.com/2644-3260/full/v1/i3/87.htm
DOI: https://dx.doi.org/10.35711/aimi.v1.i3.87

INTRODUCTION

In this new era of health-related technology and medical advances, artificial intelligence (AI) has put down roots making it possible to teach computers to do an intelligence human task, thus emerging as a problem-solving tool in data analysis and improving many aspects of clinical care[1,2].

Machine learning (ML) is a subset of AI that develops computer algorithms to make predictions or decision tasks without prior explicit programmed rules. ML algorithms use iterative static methods learning from "training" data to progressively improve the model performance over time. Based on the type of learning, ML is generally divided in (1) supervised learning, which uses labelled training data to map the expected outputs; (2) unsupervised learning, which deploys unlabelled data to learn new patterns; and (3) reinforcement learning, considered as a subfield of ML using reinforcement tools in a dynamic setting[3] (Figure 1). The supervised method is the most used ML technique in medical imaging applications and, relying on the relationship between input features and expected outcomes, the ML algorithms are grouped into three broad categories: Linear, Nonlinear and Ensemble, as described in Table 1. Furthermore, based on the data features exploited by the algorithms, ML can be applied to handcrafted features as predefined features in the data set, or to non-handcrafted features, involving raw data as part of the learning process[4]. Deep learning (DL), is a subgroup of ML techniques using non-handcrafted features and it is composed of artificial neural networks (ANN) modelled as neuron multi-layered networks allowing to automatically extract features without prior labelling and perform high-level tasks[5]. The most common ANN used in medical image analysis is based on convolutional architecture [convolutional neural networks (CNN)], consisting of hidden multi-layers that compute and filter high dimensional data to obtain the correct outputs, such as detection and characterization of tumoral lesions on imaging examinations[6].

AI-based approaches have been investigated in many fields of oncology, from imaging to histopathological and molecular diagnosis. Indeed, encouraging results have been obtained in cancer imaging, especially in screening environments. Among the different available imaging modalities, computed tomography (CT) and magnetic resonance imaging (MRI) are the most widely employed due to their prominent role in oncologic patients for staging, treatment monitoring and follow-up. Moreover, the introduction of advanced imaging techniques such as perfusion CT, MRI and MRI-diffusion-weighted imaging could provide the addition of functional over morphological data to further characterize tumor phenotype and behavior. Of note, radiology and oncology share the need for precision diagnosis and prediction models, by using cross-valuable multiple parameters from medical images and clinical data. The current applications of AI in cancer imaging include the optimization of the clinical-radiological workflow (patient screening, image acquisition) and also more specific image-based tasks (cancer detection, characterization, and treatment monitoring).

In the next sections we introduce the possible applications of AI in oncology imaging (Table 2).

Clinical-radiological workflow empowerment

AI techniques can enable the aggregation of clinical and imaging data to improve screening programs' efficiency, due to the possibility to analyse a large volume of different types of data including clinical risk factors, genetic data, and imaging

Table 1 Most commonly adopted alghoritms in supervised machine learning

ML technique	ML alghoritms	Description
Linear	(1) Linear regression; and (2) Logistic regression	Linear methods are used to modelling the relationship between the dependent variable and one or more independent variables
Nonlinear	(1) Naive Bayes; (2) Decision tree; (3) k-Nearest Neighbors; (4) Support vector machines; and (5) Neural network	Nonlinear approaches are used to produce predictive insights depending on nonlinear relationships in experimental data
Ensemble	(1) Random forest; (2) Bootstrap aggregation; and (3) Stacked generalization	Ensemble techniques stack multiple models in order to improve prediction robustness and provide more accurate predictions than any individual model

ML: Machine learning.

Table 2 Possible clinical applications of artificial intelligence in oncological imaging

Clinical application	Oncologic field	Imaging modality	AI technique
Clinical-radiological workflow	Breast cancer[9]	Mammography	ML
	Image acquisition[10,11]	CT, MRI	DL
Cancer detection	Breast cancer[12,13]	Mammography	DL
	Lung cancer[14]	X-Ray, CT	DL
Tumor segmentation	Breast Cancer[17,18]	MRI	DL
Tumor characterization	Adrenal cancer[20]	MRI	ML
	Renal cancer[21]	MRI	ML
	Lung cancer[22]	CT	ML
Tumor staging	Head and neck cancer[23]	CT	ML
	Endometrial cancer[24]	MRI	ML
Treatment monitoring	Breast cancer [26]	MRI	ML

CT: Computed tomography; MRI: Magnetic resonance imaging; AI: Artificial intelligence; ML: Machine learning; DL: Deep learning.

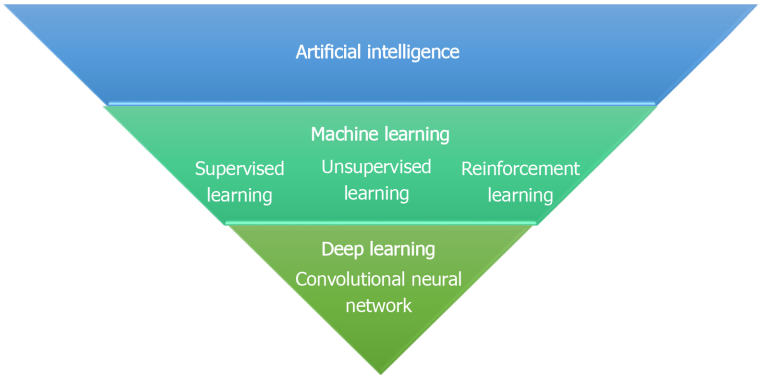

Figure 1 Diagram showing different categories of artificial intelligence.

examinations. Breast cancer surely represents a leading area of AI development, in particular in screening practices as demonstrated by recent studies that explored the impact on clinical practice of ML model in identifying individuals at increased risk of breast cancer[7,8]. Indeed, a recent study of Ming *et al*[9] investigated the performance of different ML-based techniques in predicting breast cancer risk using clinical and genetic risk factors in comparison to the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) risk prediction model; decision ML-based models yielded better results in classifying cancer from non-cancer cases and increased the predictive accuracy by 20%-25% including equal risk factors

used in the BOADICEA model. Moreover, considerable differences between the BOADICEA and risk-based ML models were observed in terms of classification for mammography surveillance according to the Swiss Surveillance Protocol, confirming the feasibility of ML prediction models in the clinical-imaging decision workup[9].

Regarding imaging acquisition and pre-processing, DL methods have shown an important impact on the reduction of radiation dose in CT examinations[10] and have been used for improving magnetic resonance imaging quality with the potential to decrease acquisition time[11].

Cancer detection

Recent evidences of AI applications include breast cancer detection in mammography, tomosynthesis, and MRI as well as identification of CT lung nodes, brain tumors, and prostate cancer on MRI.

Among these, mammographic detection of breast cancer represents a challenging image analysis task because breast cancer could be masked by healthy breast tissue. In a recent study[12], a DL AI system provided by Google Health company outperformed the radiologists involved in the mammographic screening from multiple centres in the United Kingdom (UK) and United States (US). The AI system yielded absolute reductions of 1.2% and 2.7% in false-positive and false-negative rates, respectively, in the UK test set and 5.7% and 9.4% in the US dataset. Moreover, the AI system exceeded the average performance of six expert radiologists who interpreted a sample of 500 randomly selected cases in a controlled study[12]. Similarly, Rodríguez-Ruiz *et al*[13] demonstrated that radiologists improved their diagnostic performance in detecting breast cancer on screening mammography examinations with the use of a DL-based AI system. In detail, the authors observed that radiologists improved their average area under the receiver operating characteristic curve (AUC) from 0.87 to 0.89 (P = 0.002).

In the field of lung cancer, recent research showed that a DL automatic detection algorithm achieved higher performance than the radiologist group in the detection of malignant pulmonary nodules on chest radiographs; moreover, radiologists' performance improved when DL algorithm was used as a second reader[14].

Tumor segmentation, characterization and staging

Segmentation represents one of the most challenging tasks of oncological image analysis and AI algorithms have allowed the development of systems that can enable automatic tumor segmentation. Recently, DL networks, such as CNN, have been applied in segmentation tasks gaining accurate results regarding radiotherapy treatment planning, volume measurements and, monitoring disease progression[15,16]. Indeed, increasing evidence in recent literature has highlighted the high performance of DL models in performing fully automated whole-breast segmentation to obtain reliable and robust methods for quantitative imaging analysis[17,18].

High-performance levels of AI algorithms in handling multi-dimensional data have allowed extracting and analyzing radiomic biomarkers reflecting image tumor heterogeneity thus empowering precision diagnosis and staging in cancer imaging[19] (Figure 2). For instance, with our research group, we used a combined model of MRI radiomic features and ML analysis to differentiate typical and atypical adenomas from non-adenoma adrenal lesions, which showed a better performance than the radiologist assessment[20]. Further, our group assessed the usefulness of an ML-based radiomic approach applied to MR imaging to differentiate high- from low-grade clear cell renal cell carcinoma achieving accuracy greater than 90%[21]. Reliable results and robust evidence have been providing in lung cancer diagnosis as showed in a recent work of Beig *et al*[22], which proposed a radiomic-based ML algorithm using non-contrast lung CT to distinguish non-small cell lung cancer adenocarcinomas from benign granuloma, resulting with outperforming results of AI system in comparison to the radiologists' evaluation (accuracy = 75% *vs* 61%).

Staging represents a crucial point of the oncological workflow to delineate the most appropriate treatment in a personalized and precision way. In this view, recent pilot studies have been carried out on the staging of primary tumor size, lymph nodes involvement, and distant metastasis[23]. For example, our group investigated the clinical feasibility of a combined approach of radiomics and ML-based on MR images for the identification of deep myometrial invasion in endometrial cancer in a clinical context; indeed, the integration of the developed ML algorithm improved radiologist accuracy from 82% to 100%[24].

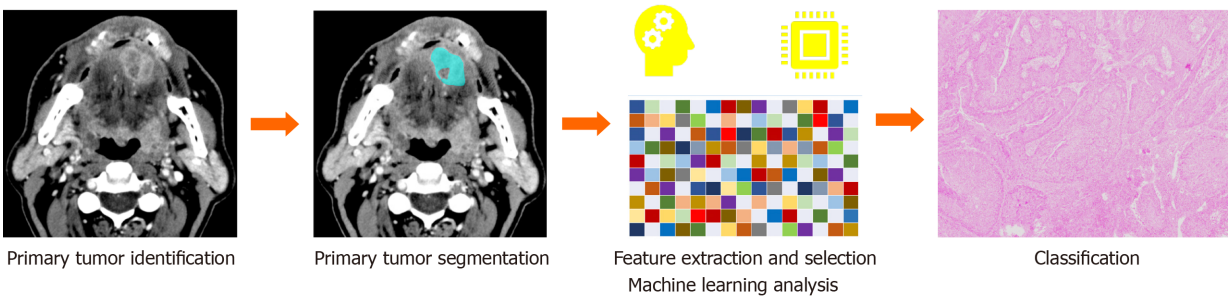

Figure 2 A schematic diagram of a radiomic and machine learning workflow pipeline applied to contrast-enhanced computed tomography images.

Treatment monitoring

Temporal follow-up of tumors as wells as the treatment response are active fields of research of AI technology to find accurate models for evaluation of efficient anticancer therapies that increase the progression-free survival of patients. In this regard, excellent results have been obtained using AI radiomics MRI-based models in predicting survival and recurrence-free survival in breast cancer[25]. Moreover, the development of AI models has been explored in breast cancer imaging to assess predictive image-based phenotypes for precision medicine, in particular to predict the response to neoadjuvant chemotherapy (NAC). In a recent study, Sutton *et al*[26] explored the usefulness of a combined radiomics MRI-based and molecular subtype-based ML model in assessing the complete pathological response (pCR) to NAC; their AI model accurately predicted pCR on MRI with an AUC of 0.88 and showed that the performance in predicting pCR increased when radiomics features were combined with molecular subtype in comparison of the solely molecular subtype results[26].

CHALLENGES AND FUTURE DIRECTIONS

AI techniques still have to face some issues to be incorporated in clinical practice. Of note, large datasets containing annotated images are needed for training of DL algorithms, but standardized imaging workflow lacks[15]. Complex AI functions are not easily interpreted by healthcare providers, and this "black box" nature could affect the acceptance of AI programs, also from the ethical and legal points of view[27]. Moreover, variability across multi-center and multi-vendor should be addressed with future studies sharing more reliable and robust validation[28].

Despite these drawbacks, AI incorporation into cancer imaging has been boosting the shift of oncology towards a precision diagnostics and personalized cancer treatment. Indeed, as previously discussed, recent literature evidence pointed out the emerging role of AI in supporting all cancer imaging pathways from screening programs to diagnostic and prognostic tasks, offering new methods to increase radiologists' performance in order to improve oncological care environment. Furthermore, the integration of AI into molecular pathological epidemiology can aid in developing pathologic signatures to stratify patients' risk and predict biological tumor behavior by using clinical, radiological and pathological data[29]. In the future, collaborations between different expertise figures, such as oncologists, epidemiologists, radiologists, pathologists and data scientists, should be encouraged to achieve harmonized and integrated development of AI systems in cancer research.

CONCLUSION

We have provided an overview of AI integration in cancer imaging, focusing on the basics of this disruptive technology and on the significant results obtained in improving clinical and radiological workup for oncological patients. Although, more evidence is still demanding the outlook of AI in cancer imaging remains bright.

REFERENCES

1 **Lakhani P**, Prater AB, Hutson RK, Andriole KP, Dreyer KJ, Morey J, Prevedello LM, Clark TJ, Geis JR, Itri JN, Hawkins CM. Machine Learning in Radiology: Applications Beyond Image Interpretation. *J Am Coll Radiol* 2018; **15**: 350-359 [PMID: 29158061 DOI: 10.1016/j.jacr.2017.09.044]

2 **Verde F**, Stanzione A, Romeo V, Cuocolo R, Maurea S, Brunetti A. Could Blockchain Technology Empower Patients, Improve Education, and Boost Research in Radiology Departments? *J Digit Imaging* 2019; **32**: 1112-1115 [PMID: 31197561 DOI: 10.1007/s10278-019-00246-8]

3 **Choy G**, Khalilzadeh O, Michalski M, Do S, Samir AE, Pianykh OS, Geis JR, Pandharipande PV, Brink JA, Dreyer KJ. Current Applications and Future Impact of Machine Learning in Radiology. *Radiology* 2018; **288**: 318-328 [PMID: 29944078 DOI: 10.1148/radiol.2018171820]

4 **Goldenberg SL**, Nir G, Salcudean SE. A new era: artificial intelligence and machine learning in prostate cancer. *Nat Rev Urol* 2019; **16**: 391-403 [PMID: 31092914 DOI: 10.1038/s41585-019-0193-3]

5 **Yang YJ**, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019; **25**: 1666-1683 [PMID: 31011253 DOI: 10.3748/wjg.v25.i14.1666]

6 **Chartrand G**, Cheng PM, Vorontsov E, Drozdzal M, Turcotte S, Pal CJ, Kadoury S, Tang A. Deep Learning: A Primer for Radiologists. *Radiographics* 2017; **37**: 2113-2131 [PMID: 29131760 DOI: 10.1148/rg.2017170077]

7 **Sheth D**, Giger ML. Artificial intelligence in the interpretation of breast cancer on MRI. *J Magn Reson Imaging* 2020; **51**: 1310-1324 [PMID: 31343790 DOI: 10.1002/jmri.26878]

8 **Ming C**, Viassolo V, Probst-Hensch N, Chappuis PO, Dinov ID, Katapodi MC. Machine learning techniques for personalized breast cancer risk prediction: comparison with the BCRAT and BOADICEA models. *Breast Cancer Res* 2019; **21**: 75 [PMID: 31221197 DOI: 10.1186/s13058-019-1158-4]

9 **Ming C**, Viassolo V, Probst-Hensch N, Dinov ID, Chappuis PO, Katapodi MC. Machine learning-based lifetime breast cancer risk reclassification compared with the BOADICEA model: impact on screening recommendations. *Br J Cancer* 2020; **123**: 860-867 [PMID: 32565540 DOI: 10.1038/s41416-020-0937-0]

10 **Greffier J**, Hamard A, Pereira F, Barrau C, Pasquier H, Beregi JP, Frandon J. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. *Eur Radiol* 2020; **30**: 3951-3959 [PMID: 32100091 DOI: 10.1007/s00330-020-06724-w]

11 **Montagnon E**, Cerny M, Cadrin-ChÃªnevert A, Hamilton V, Derennes T, Ilinca A, Vandenbroucke-Menu F, Turcotte S, Kadoury S, Tang A. Deep learning workflow in radiology: a primer. *Insights Imaging* 2020; **11**: 22 [PMID: 32040647 DOI: 10.1186/s13244-019-0832-5]

12 **McKinney SM**, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, Back T, Chesus M, Corrado GC, Darzi A, Etemadi M, Garcia-Vicente F, Gilbert FJ, Halling-Brown M, Hassabis D, Jansen S, Karthikesalingam A, Kelly CJ, King D, Ledsam JR, Melnick D, Mostofi H, Peng L, Reicher JJ, Romera-Paredes B, Sidebottom R, Suleyman M, Tse D, Young KC, De Fauw J, Shetty S. International evaluation of an AI system for breast cancer screening. *Nature* 2020; **577**: 89-94 [PMID: 31894144 DOI: 10.1038/s41586-019-1799-6]

13 **RodrÃguez-Ruiz A**, Krupinski E, Mordang JJ, Schilling K, Heywang-KÃ¶brunner SH, Sechopoulos I, Mann RM. Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. *Radiology* 2019; **290**: 305-314 [PMID: 30457482 DOI: 10.1148/radiol.2018181371]

14 **Nam JG**, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, Vu TH, Sohn JH, Hwang S, Goo JM, Park CM. Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology* 2019; **290**: 218-228 [PMID: 30251934 DOI: 10.1148/radiol.2018180237]

15 **Cuocolo R**, Caruso M, Perillo T, Ugga L, Petretta M. Machine Learning in oncology: A clinical appraisal. *Cancer Lett* 2020; **481**: 55-62 [PMID: 32251707 DOI: 10.1016/j.canlet.2020.03.032]

16 **Deig CR**, Kanwar A, Thompson RF. Artificial Intelligence in Radiation Oncology. *Hematol Oncol Clin North Am* 2019; **33**: 1095-1104 [PMID: 31668208 DOI: 10.1016/j.hoc.2019.08.003]

17 **Zhang L**, Mohamed AA, Chai R, Guo Y, Zheng B, Wu S. Automated deep learning method for whole-breast segmentation in diffusion-weighted breast MRI. *J Magn Reson Imaging* 2020; **51**: 635-643 [PMID: 31301201 DOI: 10.1002/jmri.26860]

18 **Zhang Y**, Chen JH, Chang KT, Park VY, Kim MJ, Chan S, Chang P, Chow D, Luk A, Kwong T, Su MY. Automatic Breast and Fibroglandular Tissue Segmentation in Breast MRI Using Deep Learning by a Fully-Convolutional Residual Neural Network U-Net. *Acad Radiol* 2019; **26**: 1526-1535 [PMID: 30713130 DOI: 10.1016/j.acra.2019.01.012]

19 **Gitto S**, Cuocolo R, Albano D, Chianca V, Messina C, Gambino A, Ugga L, Cortese MC, Lazzara A, Ricci D, Spairani R, Zanchetta E, Luzzati A, Brunetti A, Parafioriti A, Sconfienza LM. MRI radiomics-based machine-learning classification of bone chondrosarcoma. *Eur J Radiol* 2020; **128**: 109043 [PMID: 32438261 DOI: 10.1016/j.ejrad.2020.109043]

20 **Romeo V**, Maurea S, Cuocolo R, Petretta M, Mainenti PP, Verde F, Coppola M, Dell'Aversana S, Brunetti A. Characterization of Adrenal Lesions on Unenhanced MRI Using Texture Analysis: A Machine-Learning Approach. *J Magn Reson Imaging* 2018; **48**: 198-204 [PMID: 29341325 DOI: 10.1002/jmri.25954]

21 **Stanzione A**, Ricciardi C, Cuocolo R, Romeo V, Petrone J, Sarnataro M, Mainenti PP, Improta G, De Rosa F, Insabato L, Brunetti A, Maurea S. MRI Radiomics for the Prediction of Fuhrman Grade in Clear Cell Renal Cell Carcinoma: a Machine Learning Exploratory Study. *J Digit Imaging*2020 [PMID: 32314070 DOI: 10.1007/s10278-020-00336-y]

22 **Beig N**, Khorrami M, Alilou M, Prasanna P, Braman N, Orooji M, Rakshit S, Bera K, Rajiah P, Ginsberg J, Donatelli C, Thawani R, Yang M, Jacono F, Tiwari P, Velcheti V, Gilkeson R, Linden P, Madabhushi A. Perinodular and Intranodular Radiomic Features on Lung CT Images Distinguish Adenocarcinomas from Granulomas. *Radiology* 2019; **290**: 783-792 [PMID: 30561278 DOI: 10.1148/radiol.2018180910]

23 **Romeo V**, Cuocolo R, Ricciardi C, Ugga L, Cocozza S, Verde F, Stanzione A, Napolitano V, Russo D, Improta G, Elefante A, Staibano S, Brunetti A. Prediction of Tumor Grade and Nodal Status in Oropharyngeal and Oral Cavity Squamous-cell Carcinoma Using a Radiomic Approach. *Anticancer Res*

2020; **40**: 271-280 [PMID: 31892576 DOI: 10.21873/anticanres.13949]

24 **Stanzione A**, Cuocolo R, Del Grosso R, Nardiello A, Romeo V, Travaglino A, Raffone A, Bifulco G, Zullo F, Insabato L, Maurea S, Mainenti PP. Deep Myometrial Infiltration of Endometrial Cancer on MRI: A Radiomics-Powered Machine Learning Pilot Study. *Acad Radiol* 2020 [PMID: 32229081 DOI: 10.1016/j.acra.2020.02.028]

25 **Reig B**, Heacock L, Geras KJ, Moy L. Machine learning in breast MRI. *J Magn Reson Imaging* 2020; **52**: 998-1018 [PMID: 31276247 DOI: 10.1002/jmri.26852]

26 **Sutton EJ**, Onishi N, Fehr DA, Dashevsky BZ, Sadinski M, Pinker K, Martinez DF, Brogi E, Braunstein L, Razavi P, El-Tamer M, Sacchini V, Deasy JO, Morris EA, Veeraraghavan H. A machine learning model that classifies breast cancer pathologic complete response on MRI post-neoadjuvant chemotherapy. *Breast Cancer Res* 2020; **22**: 57 [PMID: 32466777 DOI: 10.1186/s13058-020-01291-w]

27 **Carter SM**, Rogers W, Win KT, Frazer H, Richards B, Houssami N. The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. *Breast* 2020; **49**: 25-32 [PMID: 31677530 DOI: 10.1016/j.breast.2019.10.001]

28 **Thrall JH**, Li X, Li Q, Cruz C, Do S, Dreyer K, Brink J. Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success. *J Am Coll Radiol* 2018; **15**: 504-508 [PMID: 29402533 DOI: 10.1016/j.jacr.2017.12.026]

29 **Ogino S**, Nowak JA, Hamada T, Milner DA Jr, Nishihara R. Insights into Pathogenic Interactions Among Environment, Host, and Tumor at the Crossroads of Molecular Pathology and Epidemiology. *Annu Rev Pathol* 2019; **14**: 83-103 [PMID: 30125150 DOI: 10.1146/annurev-pathmechdis-012418-012818]

Artificial Intelligence in Medical Imaging

Submit a Manuscript: https://www.f6publishing.com

Artif Intell Med Imaging 2020 September 28; 1(3): 94-107

DOI: 10.35711/aimi.v1.i3.94

ISSN 2644-3260 (online)

ORIGINAL ARTICLE

Basic Study

Predicting a live birth by artificial intelligence incorporating both the blastocyst image and conventional embryo evaluation parameters

Yasunari Miyagi, Toshihiro Habara, Rei Hirata, Nobuyoshi Hayashi

ORCID number: Yasunari Miyagi 0000-0003-0962-033X; Toshihiro Habara 0000-0003-3853-8044; Rei Hirata 0000-0002-2248-4224; Nobuyoshi Hayashi 0000-0001-6576-3066.

Author contributions: Miyagi Y, Habara T, R Hirata, and Hayashi N designed and coordinated the study; Miyagi Y and Hayashi N supervised the project; Habara T, and R Hirata acquired and validated data; Miyagi Y developed artificial intelligence software, analyzed and interpreted data, and wrote draft; Hayashi N set up project administration; Miyagi Y, Habara T, R Hirata, and Hayashi N wrote the manuscript; and all authors approved the final version of the article.

Institutional review board statement: The study was reviewed and approved by the Institutional Review Board at Okayama Couples' Clinic.

Conflict-of-interest statement: The authors declare no conflict of interest.

Data sharing statement: No informed consent was not obtained for data sharing. No additional data are available.

Open-Access: This article is an

Yasunari Miyagi, Department of Artificial Intelligence, Medical Data Labo, Okayama 703-8267, Japan

Yasunari Miyagi, Department of Gynecologic Oncology, Saitama Medical University International Medical Center, Hidaka 350-1298, Saitama, Japan

Toshihiro Habara, Rei Hirata, Nobuyoshi Hayashi, Department of Reproduction, Okayama Couples' Clinic, Okayama 701-1152, Japan

Corresponding author: Yasunari Miyagi, MD, PhD, Director, Professor, Surgeon, Department of Artificial Intelligence, Medical Data Labo, 289-48 Yamasaki, Naka ward, Okayama 703-8267, Japan. ymiyagi@mac.com

Abstract

BACKGROUND

The achievement of live birth is the goal of assisted reproductive technology in reproductive medicine. When the selected blastocyst is transferred to the uterus, the degree of implantation of the blastocyst is evaluated by microscopic inspection, and the result is only about 30%-40%, and the method of predicting live birth from the blastocyst image is unknown. Live births correlate with several clinical conventional embryo evaluation parameters (CEE), such as maternal age. Therefore, it is necessary to develop artificial intelligence (AI) that combines blastocyst images and CEE to predict live births.

AIM

To develop an AI classifier for blastocyst images and CEE to predict the probability of achieving a live birth.

METHODS

A total of 5691 images of blastocysts on the fifth day after oocyte retrieval obtained from consecutive patients from January 2009 to April 2017 with fully deidentified data were retrospectively enrolled with explanations to patients and a website containing additional information with an opt-out option. We have developed a system in which the original architecture of the deep learning neural network is used to predict the probability of live birth from a blastocyst image and CEE.

RESULTS

Manuscript source: Invited manuscript

Received: August 24, 2020
Peer-review started: August 24, 2020
First decision: September 13, 2020
Revised: September 19, 2020
Accepted: September 19, 2020
Article in press: September 19, 2020
Published online: September 28, 2020

P-Reviewer: Boon CS, Hou Y, Karmazanovsky GG
S-Editor: Wang JL
L-Editor: A
P-Editor: Li JH

The live birth rate was 0.387 (= 1587/4104 cases). The number of independent clinical information for predicting live birth is 10, which significantly avoids multicollinearity. A single AI classifier is composed of ten layers of convolutional neural networks, and each elementwise layer of ten factors is developed and obtained with 42792 as the number of training data points and 0.001 as the L2 regularization value. The accuracy, sensitivity, specificity, negative predictive value, positive predictive value, Youden J index, and area under the curve values for predicting live birth are 0.743, 0.638, 0.789, 0.831, 0.573, 0.427, and 0.740, respectively. The optimal cut-off point of the receiver operator characteristic curve is 0.207.

CONCLUSION

AI classifiers have the potential of predicting live births that humans cannot predict. Artificial intelligence may make progress in assisted reproductive technology.

Key Words: Artificial intelligence; Blastocyst; Deep learning; Live birth; Machine learning; Neural network

Core Tip: The feasibility of predicting live birth by artificial intelligence (AI) combining blastocyst images and conventional embryo evaluation parameters (CEE) is investigated because there is no human method to predict live birth from blastocyst image. Deep learning of blastocyst images is performed by using the original conventional neural network, and the elementwise layer network is used for independent CEE factors to develop a single AI classifier, the accuracy, sensitivity, specificity and area under the curve values used to predict live birth by the AI are 0.743, 0.638, 0.789, and 0.740, respectively.

Citation: Miyagi Y, Habara T, Hirata R, Hayashi N. Predicting a live birth by artificial intelligence incorporating both the blastocyst image and conventional embryo evaluation parameters. *Artif Intell Med Imaging* 2020; 1(3): 94-107
URL: https://www.wjgnet.com/2644-3260/full/v1/i3/94.htm
DOI: https://dx.doi.org/10.35711/aimi.v1.i3.94

INTRODUCTION

The achievement of live birth is the goal of assisted reproductive technology in reproductive medicine. Miscarriage or embryo developmental failure can cause cost and time loss, and bring the negative psychological outcome to the patient. Although the morphological structures have been studied, the prognosis of the developmental ability of oocytes has not yet been found[1]. Time-lapse microscopy and conventional morphological evaluations recently studied are not sufficient to ensure the thriving of the embryo after transfer[2]. The feasibility of investigating time-lapse imaging has not yet been established. Preimplantation genetic testing for aneuploidy[3,4], which is an invasive procedure for embryos, is the subject of ethical considerations. Since embryos are genetically heterogeneous, the chromosomal profile of biopsy samples does not always reflect the rest of the profile[5]. After all, no method has been established in practice to use morphological analysis and/or non-morphological analysis to predict the live birth of a blastocyst.

Recently, artificial intelligence (AI) has been developed[6] and investigated as a diagnostic tool in reproductive medicine. *e.g.*, predicting the viability of embryos can lead to a sensitivity of 70.1% for viable embryos, and a specificity of 60.5% for non-viable embryos[7]. A report showed that the AI classifier was used to classify images of mature blastocysts, which appeared to be the final stage prior to freezing or transfer, and the most important embryo stage for evaluating assisted reproductive technology demonstrated the potential for predicting the probability of live birth[8]. Our report (2019) is used to apply deep learning in convolutional neural networks (CNN) to the prediction of live births[9-12] to blastocyst images classified by maternal age,

demonstrated that the accuracy, sensitivity, specificity, positive predictive value, and negative predictive value area under the curve (AUC) were 0.732, 0.673, 0.753, 0.404, 0.862 and 0.726, respectively[13]. To the best of our knowledge, these are unique values for predicting live birth through image recognition of blastocyst images. We have previously reported live birth predictions using multivariate logistic regression function in combination with conventional embryo evaluation (CEE) (*e.g.* maternal age, body mass index, *etc.*) and the application of deep learning, which was applied to blastocyst images that were also classified by age; this method was defined as a combination method[13,14] in which the accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and AUC value for all ages were 0.721, 0.779, 0.704, 0.400, 0.885 and 0.773, respectively (2019). This combination method seemed to be better than CEE.

AI can be trained through images and non-images (such as numbers[15]) simultaneously. The classifier made by AI can convert data composed of image data and non-image data into a confidence score, which is an estimated probability of belonging to a target category (such as a live birth category). Therefore, when inputting images and non-images, an AI classifier trained on image and non-image data can generate confidence scores. The AI feature that can convert images into probabilities seems to be an outstanding advantage. Compared with AI classifiers trained only by images, AI classifiers trained with more information (including image and non-image data) may show better results. Then it may be necessary to investigate whether a single AI classifier of deep learning might demonstrate better predictability than the combination method when applied to both the blastocyst image and independent CEE factors, which were not classified by age but included age, to predict a live birth. Although it is necessary to use a combination method to create multiple AI classifiers, it is not necessary to classify a single AI classifier by age as an independent factor of CEE[16,17]. Therefore, we constructed the original neural network architecture of the AI classifier as a pilot study and demonstrated the feasibility of the classifier compared with the combination method.

MATERIALS AND METHODS

Patients and data preparation

The study collected images of blastocysts with morphological features and clinical information obtained from consecutive patients at the Okayama Couples' Clinic from January 1, 2009, to April 30, 2017, with completely deidentified data were enrolled. Only elective single embryo transfer is performed. Track all blastocysts to confirm whether the result is a live birth or a non-live birth. This retrospective study was approved by the Institutional Review Board (IRB) of Okayama Couples' Clinic (IRB number 18000128-5). This non-interventional study provides patients with the option to opt-out with additional information on the clinic's website.

CEE

All blastocysts with clinical information and morphological features, such as maternal age, body mass index, past embryo transfer time, *in vitro* fertilization time, anti-Müllerian hormone value, FSH value, blastocyst grade on day 3, embryo cryopreservation day, Trophoblast grade, inner cell mass grade, number of blastomeres on the 3rd day after insemination, the average diameter of blastocysts, antral follicle count, the existence of immune sterility, the existence of oviduct infertility, the existence of endometriosis, insemination procedures, ovarian stimulation method, the grade of smooth endoplasmic reticulum cluster, degree of blastocyst expansion, presence of vacuoles, refractile body, male age, and male body mass index, were collected to evaluate the outcome of live birth *vs* non-live birth. This information was provided by doctors and embryologists engaged in clinical practice for over twenty years and who have implemented standard laboratory practices related to embryo morphological evaluations according to the 2011 international consensus meeting[18].

The relationships between each factor in CEE and live birth were assessed. Then, we obtained univariate regression functions. Significant factors without multicollinearity which indicated a state of strong correlations between variables were selected as independent factors to predict live birth.

Blastocyst images

As a routine conventional microscopic observation at magnification of 400 times, a single clear image of the blastocyst is captured at about 115 h after insemination or about 139 h if the blastocyst is less than approximate 120 μm in diameter. According to a published report[14], each image is cropped into a square and then saved in size of 50 × 50 pixels to provide the best accuracy. The picture has been de-identified so as not to identify the person.

Preparation for AI

The deidentified data set included all the factors in CEE, images of the blastocysts that resulted in miscarriages, non-live births, or live births were transferred to the AI system off-line.

AI classifier

AI classification programs were developed as shown in Figure 1. AI classifiers which were made up of both CNN[19-24] with L2 regularization[25,26] and elementwise functions that apply a function to each element of a tensor for each factor of the CEE to obtain the probability of predicting a live birth or non-live birth, as shown in Figure 1. We introduced deep learning for images with a published CNN architecture except for a softmax layer[14]. The CNN in the image consisted of 10 Layers with a combination of convolutional layers with multiple kernel sizes[27-29], pooling layers[30-33], flattened layers[34], linear layers[35,36] and rectified linear layers unit layers[37,38]. On the other hand, we also performed elementwise functions that we reported for the CEE factors[14].

Then, all tensors from the convolutional network for image and scholars by the elementwise functions for the factors of the CEE were catenated and inputted into a batch normalization layer[39]. Then, the data was placed in a linear layer and a softmax layer[40,41] which presented the probability of live birth or non-live birth.

The appropriate number of training datasets was investigated by evaluating the accuracies using the ten-fold cross-validation method[42-44]. Firstly, all data were divided into test datasets and training datasets randomly in a ratio of one to nine. Four-fifths of the training data set was used as the AI training dataset. The remainder, one-fifth, of the dataset, was defined as the validation dataset. The AI training dataset, validation dataset, and non-overlapping test dataset were created in this fashion. The AI classifier was trained by an AI training dataset with concurrent validation by the validation dataset, and then the AI classifier was evaluated with the test dataset. The training dataset is augmented by rotating images, as is often performed in the AI classifier process known as data augmentation, because the blastocyst image processing with any degree of rotation can produce images, resulting in different vector data of the same category[14]. Repeat this procedure ten times to incorporate all the data. Investigate the number of training data points until the accuracy value is the largest possible while keeping the variance of the accuracy value as small as possible. Therefore, this process can temporarily display an appropriate amount of training data to more accurately verify the prediction. Then, by varying the hyperparameters and the number of training data points, the best AI classifier showing the best accuracy was finally selected during the early stopping procedure. By comparing with the combination method, the feasibility of the new method is evaluated.

Development environment

The tools and conditions for development used are as follows: Intel Core i5 running Windows 10 (Redmond, WA, United States), 32 GB (Santa Clara, CA, United States) and NVIDIA GeForce GTX 1080 Ti (Santa Clara, CA, United States) and Wolfram Language 12.0 (Wolfram Research, Champaign, IL, United States).

Statistical analysis

Wolfram Language 12.0 is used for all statistical analyses. One-way analysis of variance test and univariate regression analysis was used. $P < 0.05$ was considered to indicate statistical significance.

RESULTS

Clinical information and morphological features

There were 5691 blastocysts, among which the outcome of live birth and non-live birth were 1587/4104, respectively. Images, morphological feature data, and clinical

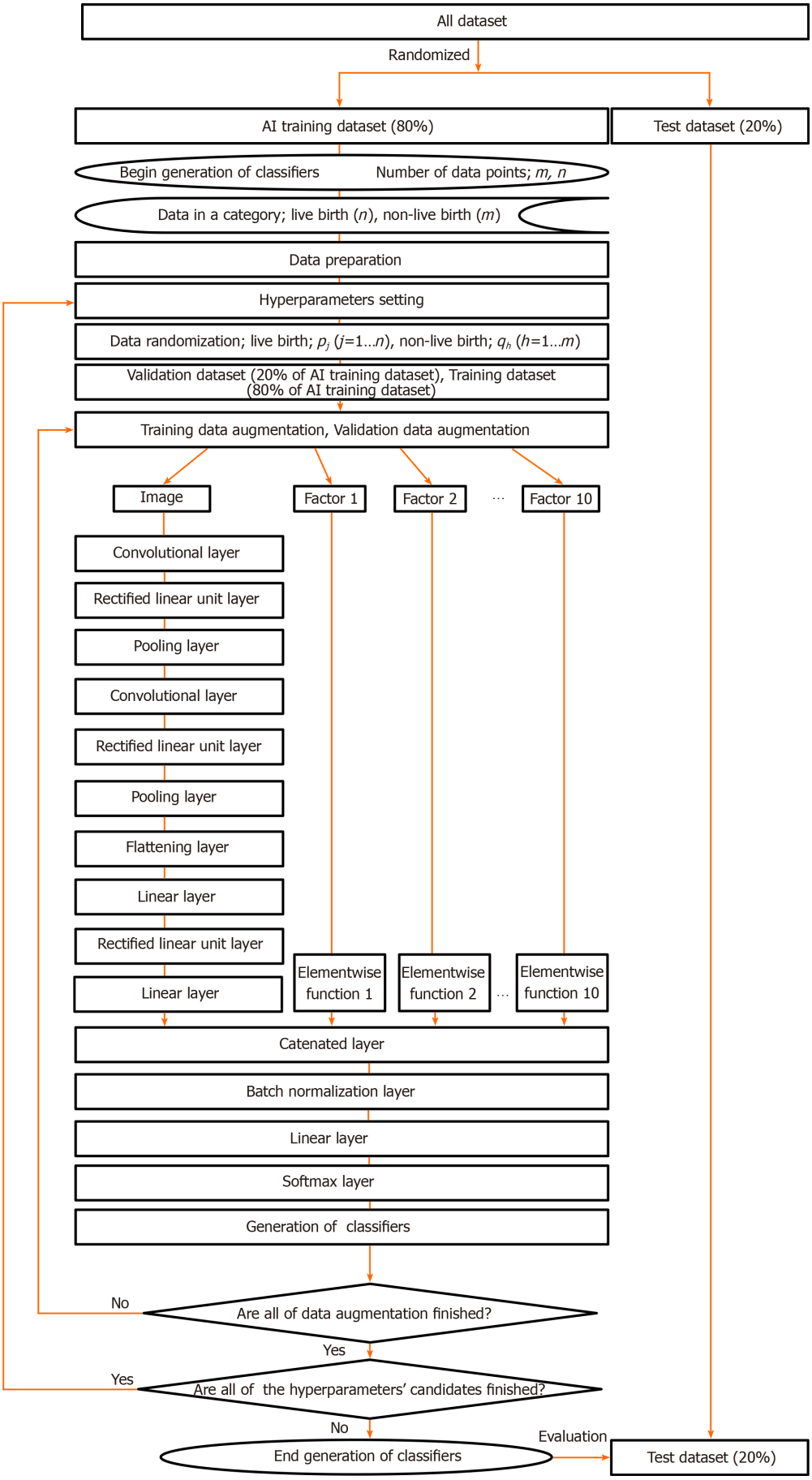

Figure 1 Flowchart for generating the artificial intelligence classifiers. The artificial intelligence classifier consisted of a combination of 10 Layers of a

convolutional neural network for an image and each elementwise layer often significant factors of the conventional embryo evaluation (CEE) that we had reported[14]. The ten factors chosen as independent factors to predict live birth were age, the number of embryo transfers, anti-Müllerian hormone concentration, day-3 blastomere number, grade on day 3, embryo cryopreservation day, inner cell mass, trophectoderm, average diameter, and body mass index. The functions in the elementwise layer for each factor of the CEE are shown as formulas in Table 1. The image processing and ten factors of the conventional embryo evaluation tensor were combined at the catenated layer. AI: Artificial intelligence.

information were obtained. The live birth rate was 0.387. Table 1 shows the independent clinical information and morphological features for predicting live births. The mean ± SD/median/range for age, number of embryo transfers, anti-Müllerian hormone concentration (ng/mL), day-3 blastomere number, grade on day 3 (class A = 1, B = 2, C = 3, D = 4), embryo cryopreservation day (day 5 = 1, day 6 = 2), inner cell mass (A = 1, B = 2, C = 3), trophectoderm (A = 1, B = 2, C = 3), average diameter (μm) and body mass index (kg/m^2) were 35.75 ± 4.78/36/20-48; 2.75 ± 2.36/2/1-30; 3.91 ± 3.54/2.94/0.0-32.2; 8.03 ± 1.74/8/2-17; 1.87 ± 0.57/2/1-4; 1.20 ± 0.40/1/1-2; 1.58 ± 0.55/2/1-3; 2.01 ± 0.75/2/1-3; 154.77 ± 24.12/153.8/81.3-242.5; and 21.30 ± 3.16/20.6/13.9-43.3, respectively.

Univariate regression functions

The univariate regression functions in order to use at the elementwise layer in the neural network for each were as follows: Age, $k/[1 + \mathrm{Exp}\ (\beta_0 + \beta_1 x)]$, $\beta_0 = -10.742$, $\beta_1 = 0.284$, $k = 0.451$; number of embryo transfers, $1/[1 + \mathrm{Exp}\ (\beta_0 + \beta_1 x)]$, $\beta_0 = 0.635$, $\beta_1 = 0.156$; anti-Müllerian hormone concentration (ng/mL), $1/[1 + \mathrm{Exp}\ (\beta_0 + \beta_1 x)]$, $\beta_0 = 1.282$, $\beta_1 = 0.062$; day-3 blastomere number, $k/(2\pi\sigma^2)^{0.5}\ \mathrm{Exp}\ [(x-m)^2/(2\sigma^2)]$, $\sigma = 4.668$, $m = 11.624$, $k = 4.643$; grade on day 3 (class A = 1, B = 2, C = 3, D = 4), $k/[1 + \mathrm{Exp}\ (\beta_0 + \beta_1 x)]$; $\beta_0 = -7.967$, $\beta_1 = 2.584$, $k = 0.319$; embryo cryopreservation day (day 5 = 1, day 6 = 2), $\beta_0 + \beta_1 x$; $\beta_0 = 0.435$, $\beta_1 = -0.131$; inner cell mass (A = 1, B = 2, C = 3), $\beta_0 + \beta_1 x$, $\beta_0 = 0.479$, $\beta_1 = -0.131$; trophectoderm (A = 1, B = 2, C = 3), $\beta_0 + \beta_1 x$; $\beta_0 = 0.526$, $\beta_1 = -0.124$; averaged diameter (μm), $1/[1 + \mathrm{Exp}\ (\beta_0 + \beta_1 x)]$, $\beta_0 = 2.623$, $\beta_1 = -0.011$; and body mass index (kg/m^2), $1/[1 + \mathrm{Exp}\ (\beta_0 + \beta_1 x)]$, $\beta_0 = -0.631$, $\beta_1 = 0.079$.

The approximate number of training data points

Overview of the accuracy profile as a function of the approximate number of training data points to study the appropriate amount of training data are shown in the left panel of Figure 2. The accuracy values were classified with L2 regularization values. High accuracies were obtained when the number of training data points was between 25605 and 45468. The mean of standard deviation of each parameter in the training data set are 0.0163, 0.0090, 0.0082, 0.0075, 0.009, 0.0121, 0.0086, and 0.0071, respectively. Although there is training data, there is no significant difference in standard deviation ($P = 0.223$ by one-way analysis of variance test). The same data were converted to a two-dimensional contour plot of accuracy as a function of the number of training data points and the number of L2-regularization values (right panel in Figure 2). The brighter area that indicated higher accuracy was observed when the number of the training data points was between 25605 and 45468 and when the L2-regularization values were less than 0.1.

AI classifier

Therefore, the best AI classifier is investigated. When the number of training data points is between 25605 and 45468 and the L2-regularization values are less than 0.1, the best AI classifier will exist. Finally, the best AI classifier was obtained with 42792 training data points and 0.001 L2 regularization values. The accuracy, sensitivity, specificity, positive predictive value, negative predictive value, Youden J index[45] and area under the curve (mean ± SE) obtained by the AI classifier are 0.743, 0.638, 0.789, 0.573, 0.831, 0.427 and 0.740 ± 0.031, respectively, as shown in Table 2. The optimal cut-off point of the receiver operator characteristic (ROC) curve[46] is 0.207. The classification time per case is less than 0.2 s.

DISCUSSION

Here, a single AI classifier for deep learning with CNN using blastocyst images and elementwise layers using independent factors of the morphological features and clinical information of the CEE is developed. When the patient's age is less than 39 years old, this integrated AI classifier is superior to the combination method in terms

Table 1 The morphological features and clinical information of 5691 blastocysts and the univariate regression formulas[13] of the independent factors for predicting the probability of live birth

Independent factors	mean ± SD	Median	Minimum	Maximum	Formulas	Coefficients
Age	35.75 ± 4.78	36	20	48	$k/[1 + Exp (\beta_0 + \beta_1 x)]$	β_0 = -10.742 ± 4.106 (P = 0.0089); β_1 = 0.284 ± 0.109 (P = 0.0088); K = 0.451
Number of embryo transfers procedures in the past	2.75 ± 2.36	2	1	30	$1/[1 + Exp (\beta_0 + \beta_1 x)]$	β_0 = 0.635 ± 1.158 (P = 0.584); β_1 = 0.156 ± 0.123 (P = 0.204)
Anti-Müllerian hormone concentration (ng/mL)	3.91 ± 3.54	2.94	0.0	32.2	$1/[1 + Exp (\beta_0 + \beta_1 x)]$	β_0 = 1.282 ± 2.640 (P = 0.627); β_1 = 0.062 ± 0.139 (P = 0.678)
Day-3 blastomere number	8.03 ± 1.74	8	2	17	$k/(2\pi\sigma^2)^{1/2} Exp (-(x-m)^2/(2\sigma^2))$	σ = 4.668 ± 0.773 (P = 4.179 × 10^{-5}); m = 11.624 ± 0.663 (P = 1.969 × 10^{-10}); K = 4.643 ± 0.611 (P = 3.91 × 10^{-6})
Grade on day 3 (Class A = 1, B = 2, C = 3, D = 4)	1.87 ± 0.57	2	1	4	$k/[1 + Exp (\beta_0 + \beta_1 x)]$	β_0 = -7.967 ± 8.012 (P = 0.320); β_1 = 2.584 ± 2.582 (P = 0.317); K = 0.319
Embryo cryopreservation day (Day 5 = 1, Day 6 = 2)	1.20 ± 0.40	1	1	2	$\beta_0 + \beta_1 x$	β_0 = 0.435; β_1 = -0.131
Inner cell mass (A = 1, B = 2, C = 3)	1.58 ± 0.55	2	1	3	$\beta_0 + \beta_1 x$	β_0 = 0.479 ± 0.037 (P = 0.049); β_1 = -0.131 ± 0.017 (P = 0.083)
Trophectoderm (A = 1, B = 2, C = 3)	2.01 ± 0.75	2	1	3	$\beta_0 + \beta_1 x$	β_0 = 0.526 ± 0.002 (P = 0.0026); β_1 = -0.124 ± 0.001 (P = 0.005)
Average diameter (μm)	154.77 ± 24.12	153.8	81.3	242.5	$1/[1 + Exp (\beta_0 + \beta_1 x)]$	β_0 = 2.623± 5.312 (P = 0.621); β_1 = -0.011 ± 0.030 (P = 0.723)
Body mass index (kg/m^2)	21.30 ± 3.16	20.6	13.9	43.3	$1/[1 + Exp (\beta_0 + \beta_1 x)]$	β_0 = -0.631± 0.844 (P = 0.454); β_1 = 0.079 ± 0.035 (P = 0.026)

Each formula was determined to fit the data distribution. Coefficients are shown as the mean ± SE.

Table 2 Discrimination ability of the best classifier of the original neural network architecture comparing the combination method[13]

Patient age (yr)	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC	95%CI of the AUC	Cut-point
AI in this study								
All ages	0.743	0.638	0.789	0.573	0.831	0.740	0.681-0.801	0.207
The combination method[13]								
All ages	0.721	0.779	0.704	0.400	0.885	0.773	0.655-0.888	0.213
< 35	0.616	0.652	0.592	0.515	0.719	0.655	0.600-0.707	0.388
35-37	0.671	0.786	0.612	0.508	0.849	0.723	0.653-0.793	0.281
38-39	0.732	0.758	0.725	0.455	0.908	0.791	0.693-0.889	0.219
40-41	0.801	0.700	0.816	0.350	0.950	0.806	0.687-0.925	0.142
≥ 42	0.784	1.000	0.773	0.171	1.000	0.888	0.713-1.063	0.037

The number of the training data was 42792 consisted of both a blastocyst image and ten factors of conventional embryo evaluation (CEE). The value of the combination method was a function of multivariate logistic regression with CEE and artificial intelligence for predicting probability for live birth from an image of the blastocyst in patients categorized by age. The optimal cut-point of live birth was the value corresponding to the point with the lowest distance to the upper-left corner of the receiver operator characteristic curve[46]. The accuracies, sensitivities, and specificities were obtained by using cut-points. PPV: Positive predictive value; NPV: Negative predictive value; AUC: Area under the time concentration curve.

of accuracy in predicting which embryo to transfer to obtain a live birth.

The accuracy value of predicting live birth is 0.743. We have previously reported that the accuracy values of predicting live birth through the CEE/AI/combined method are 0.631/0.647/0.616, 0.687/0.675/0.671, 0.725/0.697/0.732, 0.714/0.776/0.801 and 0.910/0.866/0.784 for the age categories of < 35, 35-37, 38-39, 40-41 and ≥ 42 years, respectively[13]. Our report provides 0.721 ± 0.077 (mean ± SD) as

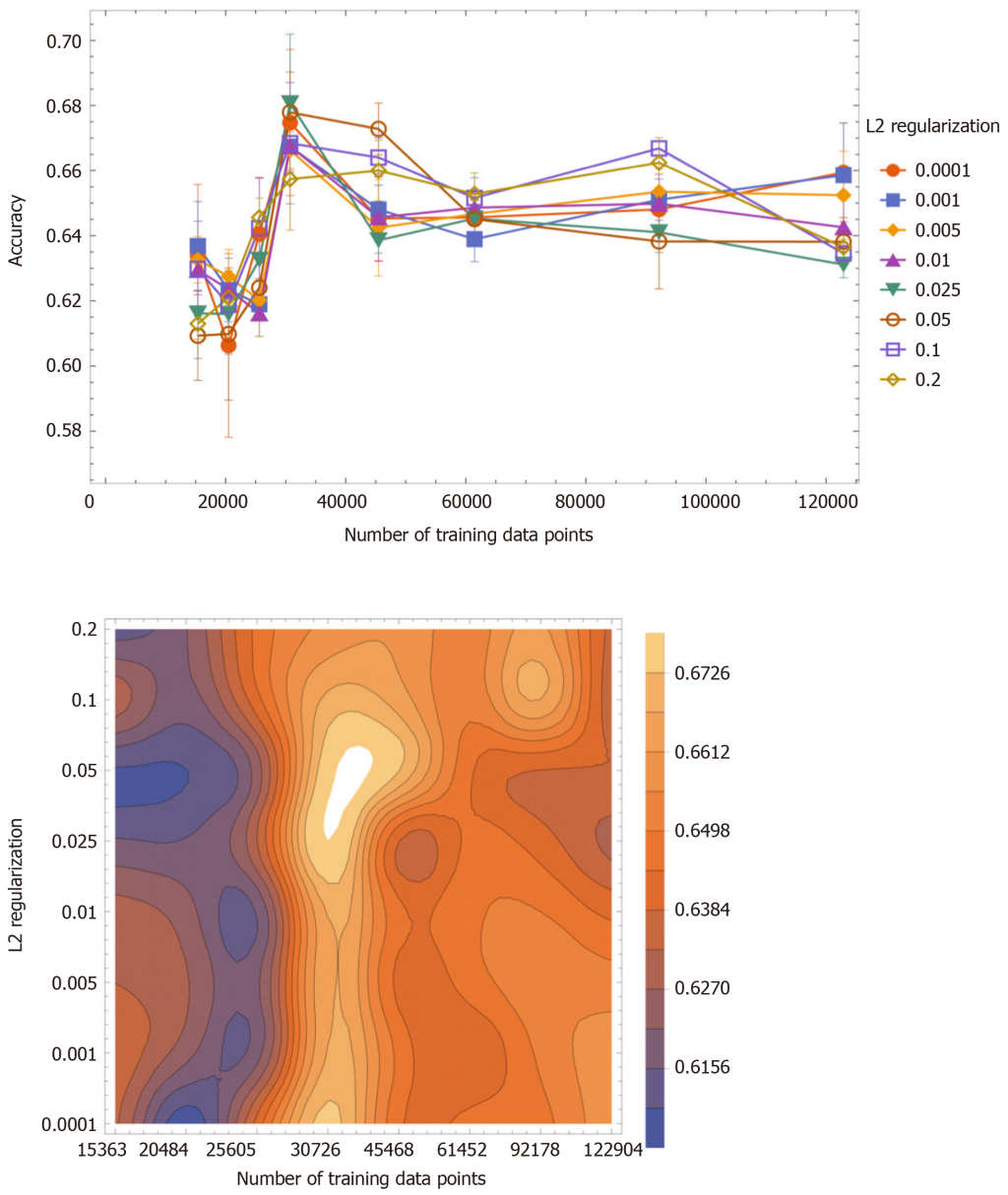

Figure 2 The accuracy value (mean ± SD) as a function of the number of training data points. The accuracy values were classified with L2 regularization values: 0.0001, 0.001, 0.005, 0.01, 0.025, 0.05, 0.1 and 0.2. High accuracy values were obtained when the number of training data was between 25605 and 45468 (left panel). The two-dimensional contour plot of the accuracy value as a function of the training data and L2 regularization values (right panel). The brighter area indicates higher accuracy. High accuracy values were observed when the number of the training data was between 25605 and 45468 and when L2 regularization values were less than 0.1.

the accuracy value. The combination method is the multivariate logistic function, the CEE probability generated by the multivariate logistic regression, and the confidence score generated by AI and the deep learning of CNN independently. However, the AI classifier in this study used both blastocyst images and CEE factors including age. These two different methods used the same data set composed of blastocyst images and CEE factors, and there was no significant difference in accuracy (P = 0.52). Therefore, in this study, the accuracy value of 0.743 as a predictor of live birth seems to be close to the average accuracy of the combination method. In this study, the results on the accuracy value are superior to the combination method that was classified by maternal age (when the patient's age is less than 39 years old); when the patient's age is greater than 39 years old, the classification method is inferior. Regarding the accuracy value, if the AI classifier in this study is used, it will be better for patients younger than 39 years old, as shown in Table 2.

Although there is no other way to predict live births, compared with AI in other medical classifiers, this single AI classifier does not seem to be good enough. The accuracy value of the AI classifier has been published, and were 0.997 for the breast cancer diagnosis[47]; 0.83-0.90 for the early diagnosis of Alzheimer's disease[48]; 0.83 for

urological dysfunctions[49]; 0.72[50], 0.50[51], 0.823[52] and 0.941[53] for colposcopy diagnosis; 0.83 for the orthopedic trauma diagnosis[54]; and 0.98 for the morphological quality of blastocysts with the evaluation by the embryologist[55]. In one report, due to the probability of live births, images of embryos classified as poor and good were scored 0.509 and 0.614, respectively[55]. The AI classifier fails to notice clinical obstacles to achieving delivery, such as uterine factors[56] (*e.g.*, uterine leiomyoma[57] and endometrial polyps[58]), endometriosis[59], ovarian function[60], oviduct obstruction[61,62], immune disorders[63,64] and the uterine microbiota[65,66]), so the prediction of blastocyst outcome through images can never reach 100%. Therefore, in this study, using 0.743 as the accuracy value for predicting live birth, as the application of AI in medicine, seems to be a moderately good result.

The values of AUC, sensitivity, and specificity are the most important statistical data for evaluating binary classification test methods because these values are independent of the distribution of the patient. The AUC in this study was 0.740 ± 0.031 (mean ± SE). In our report, the AUC values of predicted live births done by CEE/AI/combination methods are 0.651/0.634/0.655, 0.697/0.688/0.723, 0.771/0.728/0.791, 0.788/0.743/0.806 and 0.820/0.837/0.888 for the age categories of < 35, 35-37, 38-39, 40-41, and ≥ 42 years, respectively[13].

The reported AUC of the combination method is 0.773 ± 0.088 (mean ± SD). There was no significant difference between the AUC value of the AI classifier and the average AUC value of the combination method (P = 0.41). However, in this study, when the patient is younger than 37 years old, the results of the AUC value may be superior to the combination method, and when the patient is older than 37 years old, the results are inferior. Regarding AUC, if the AI classifier in this study is used, it will be better for patients younger than 37 years old. The published AUC value of the AI classifier with deep learning is 0.66, which can predict live birth[13]; 0.65 predicts live birth without aneuploidy[8]; 0.74 classifies embryos into three categories[64]; 0.826[52] for colposcopy by image, and 0.941[53] for colposcopy by image combined with HPV. Therefore, in this study, as an AI application in medicine, an AUC value of 0.740 seems to be a moderately good result.

The sensitivity of this study is 0.638. In our report, the sensitivity of CEE/AI/combination method to age category is 0.580/0.530/0.652, 0.714/0.655/0.786, 0.727/0.697/0.758, 0.700/0.650/0.700, and 0.667/0.833/1.000 for the age categories of < 35, 35-37, 38-39, 40-41, and ≥ 42 years, respectively[13]. The sensitivity of this combination method is 0.779 ± 0.134 (mean ± SD). The sensitivity of this study is inferior to the combination method (P < 0.019) and lower than the combination method of any age category.

The specificity of this study was 0.789. In our report, the specificities of CEE/AI/combination methods are 0.665/0.724/0.592, 0.673/0.685/0.612, 0.725/0.697/0.725, 0.716/0.794/0.816, and 0.922/0.867/0.773 for the age categories of < 35, 35-37, 38-39, 40-41, and ≥ 42 years, respectively[13]. The specificity of the combination method in the report is 0.704 ± 0.098 (mean ± SD). Although there is no significant difference, the specificity of this study is superior to the combination method (P = 0.052). Except for 40-41 years old, the specificity of the combination method of any age category is higher.

In this study, Youden's J index[45] was 0.427. Youden's J index (sensitivity plus specificity -1) is a statistical value that is very valuable for dichotomous diagnostic tests, and can sometimes be used for ROC analysis. In our report, the Youden's J index values of the CEE/AI/combination methods are 0.245/0.254/0.244, 0.387/0.340/0.398, 0.452/0.394/0.483, 0.416/0.444/0.516, and 0.589/0.700/0.773 for the age categories of < 35, 35-37, 38-39, 40-41, and ≥ 42 years, respectively[13]. The combination method in the report yielded 0.483 ± 0.193 (mean ± SD). There is no significant difference in Youden's J index (P = 0.519). However, in this study, the results of the Youden's J index may be superior to the combination method when the patient is younger than 37 years old, and inferior to the combination method when the patient is older than 37 years old. Regarding Youden's J index, if the AI classifier in this study is to be used, it will be better for patients younger than 37 years old. A report has been published on the Youden's J index value of the medical AI classifier. The index of LSIL/HSIL diagnosed by deep learning colposcopy is 0.682[52] and 0.789[53], while the index for predicting live birth is 0.30 without aneuploidy[8].

As for the accidental evaluation and comparison of AI only used for blastocyst images[13], its accuracy, sensitivity, specificity, positive predictive value, negative predictive value, Youden's J index, and AUC value are 0.732/0.721/0.743, 0.673/0.779/0.638, 0.753/0.704/0.789, 0.404/0.400/0.831, 0.862/0.885/0.573, 0.426/0.482/0.427 and 0.726/0.773/0.740 by AI for blastocyst image only/AI as the combination method/AI in this study. We hope that the AI or combination method in

this study will be better than other methods, but there seems to be no outstanding AI. Although the AI in this study seemed to demonstrate tendency that was superior for specificity and positive predictive value, and inferior for sensitivity and negative predictive value and AI for blastocyst image only seemed to be superior to AI in this study for negative predictive value and the value of the AUC, there was no significant superiority among the three types of AI classifiers. This result might suggest that medical imaging as the morphological features of the blastocyst was the most important among the significant parameters in the dataset. Non-image parameters may not contribute much to predicting live birth, but there is one thing. Therefore, multi-image data sets such as time-lapse photography of AI may be good candidates for predicting live births in the future. In this study, only a single image was evaluated, but it is known through time-lapse image analysis that morphology is not a static parameter, so it will be evaluated by multiple images in the future. In further research, the application of artificial intelligence consists of different neural network architectures that can process images and non-images in multiple time series at the same time, which may be better applied to time-lapse evaluation, because it has not yet shown the established method of predicting live birth. Without the intervention of more complex statistical methods, or preferably by AI applications, it may be difficult to analyze multiple data composed of images and non-images.

Not only is there no gold standard method for making neural network architectures for general targets but also images. However, the following neural networks for general image recognition have made progress: LeNet[68] (in 1998), AlexNet[40] (in 2012), GoogLeNet[36] (in 2014), ResNet[69] (in 2015), and Squeeze-and-Excitation networks[70] (in 2017). In this study, the image of the blastocyst was analyzed using CNN, and the scholar data of CEE was converted at the elementwise layers that have a function for each factor. After connecting these outputs, by using several network layers, the probability of live birth or non-live birth can be generated at the end of the neural network through the softmax function. The AI of the neural network can evaluate not only images but also non-image data. Compared with traditional statistics, this function of artificial intelligence seems to be advanced. When evaluating images through traditional statistical data, humans should define some image features (such as morphological shape and hue) before analysis, and then extract and quantitatively convert them into tensor data. Although the criteria for extracting certain features from images are indispensable for using traditional statistical data, the universality of the definition cannot be proven. On the other hand, artificial intelligence can evaluate images without any standard to extract certain features. Therefore, it is expected that AI can predict live birth through blastocyst images and CEE factors. As far as we know, there are no reports about the simultaneous use of image and non-image data for live birth prediction.

The ability of the AI classifier neural network, which consists of the CNN for the image and elementwise functions for the scholarly values of the CEE factors in this study, was almost similar to that of the published combination method[13]. The AI classifier in this study demonstrated insignificant superiority in terms of specificity, significant inferiority in terms of sensitivity, and similarity in terms of accuracy, Youden J index, and AUC. However, comparing the accuracy of the AI in this study and the combination method in patients aged 35 to 37 years in a validation study, the required sample size would be 9497000 with 0.05 and 0.20 alpha and beta errors, respectively. Also, for the AI classifier of the neural network composed of CNN for the image and some networks for the non-image data, modifications in the network architecture, hyperparameters, and an increase in the number of datasets are expected. Although further prospective studies may be required, this AI model appeared to have the potential for clinical applicability. Also, this AI classifier was a single classifier that could be easier to improve in the future, although the five AI classifier combination method would be more difficult to improve because it would require a data set for each age category, resulting in a higher number of data sets would lead. The AI classifier can display the ranking of the blastocysts to predict a live birth with decimal places, and it helps embryologists and clinicians select the blastocyst for embryo transfer. There can be a quick diagnosis of the prediction over a distance without expensive equipment when the image and CEE parameters are transmitted over the internet.

Since there is theoretically an infinite number of probabilities for the construction of the neural network architecture and numerous combinations of statistical functions, further investigations for patients are worthwhile. By selecting the hyperparameters and setting the random seed value within the program in various ways, the result can be changed, *e.g.*, the prediction accuracy can be a little better or a little worse. Similar statements can be made about the dataset. If one uses the same deep neural network

architecture and a different training dataset, for example, provided by a different institute, the prediction accuracy differs. This is one of the aspects of current AI technology. The AI in this study had not been tested for external data as an institutional joint research to validate the generalization ability. In the field of AI technology, a critical statistical method for evaluating the relative superiority of predictive ability between two classifiers is not well established. Therefore, at least the clear superiority in prediction accuracy, the advantages of the network architecture, or a wider variety of datasets that were included in the analysis should be considered before conventional practical use. To improve the AI classifier in the future and to examine not only conventional static values such as accuracy, but also robustness, stability, and reliability. It may be necessary to use the multiple-image data of blastocytes obtained by time-lapse methods from other institutes to evaluate the prediction accuracy, the increased amount of data, novel non-image data that would be significant and not yet discovered, such as genetic information or some patient biomarkers in the future, and the incorporation of the vastly improved neural network architecture.

CONCLUSION

Deep learning with a CNN for a blastocyst image and with networks of elementwise layers for independent CEE factors were used to develop the single AI classifier for predicting the probability of live birth. Due to the development of AI that does not harm the embryo, the embryo can be transferred after making the prediction. AI could bring benefits to the advancement of assisted reproductive technology.

ARTICLE HIGHLIGHTS

Research background

To acquire live births is the goal of assisted reproductive technology. No method has been established in practice to use non-morphological analysis and/or morphological analysis such as conventional morphological evaluations and time-lapse microscopy to predict the live birth of a blastocyst.

Research motivation

Artificial intelligence (AI) classifiers for blastocyst images to predict the live birth has been introduced in reproductive medicine recently.

Research objectives

The present study aimed to develop an AI classifier that combines blastocyst images and the morphological features and clinical information of the conventional embryo evaluation parameters such as maternal age to predict the probability of achieving a live birth.

Research methods

A total of 5691 images of blastocysts combined with conventional embryo evaluation parameters were used. A system in which the original architecture of the deep learning neural network was developed to predict the probability of live birth.

Research results

The number of independent clinical information for predicting live birth is 10. The best single AI classifier composed of ten layers of convolutional neural networks and each elementwise layer of ten factors was developed and obtained with 42792 as the number of training data points and 0.001 as the L2 regularization value. The accuracy, sensitivity, specificity, negative predictive value, positive predictive value, Youden J index, and area under the curve values for predicting live birth were 0.743, 0.638, 0.789, 0.831, 0.573, 0.427, and 0.740, respectively.

Research conclusions

AI classifiers have the potential of predicting live births that humans cannot predict. AI that can be trained by both morphological and non- morphological information may make progress in assisted reproductive technology.

Research perspectives

Due to the development of AI that does not harm the embryo, the embryo can be transferred after making the prediction. AI could bring benefits to the advancement of assisted reproductive technology.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the anonymous reviewers for their useful comments and suggestions on how to improve the quality of this paper.

REFERENCES

1 **Rienzi L**, Vajta G, Ubaldi F. Predictive value of oocyte morphology in human IVF: a systematic review of the literature. *Hum Reprod Update* 2011; **17**: 34-45 [PMID: 20639518 DOI: 10.1093/humupd/dmq029]
2 **Kirkegaard K**, Ahlström A, Ingerslev HJ, Hardarson T. Choosing the best embryo by time lapse versus standard morphology. *Fertil Steril* 2015; **103**: 323-332 [PMID: 25527231 DOI: 10.1016/j.fertnstert.2014.11.003]
3 **Dahdouh EM**, Balayla J, Audibert F; Genetics Committee; Wilson RD; Audibert F; Brock JA; Campagnolo C; Carroll J; Chong K; Gagnon A; Johnson JA; MacDonald W; Okun N; Pastuck M; Vallée-Pouliot K. Technical Update: Preimplantation Genetic Diagnosis and Screening. *J Obstet Gynaecol Can* 2015; **37**: 451-463 [PMID: 26168107 DOI: 10.1016/s1701-2163(15)30261-9]
4 **Brezina PR**, Kutteh WH. Clinical applications of preimplantation genetic testing. *BMJ* 2015; **350**: g7611 [PMID: 25697663 DOI: 10.1136/bmj.g7611]
5 **Gleicher N**, Metzger J, Croft G, Kushnir VA, Albertini DF, Barad DH. A single trophectoderm biopsy at blastocyst stage is mathematically unable to determine embryo ploidy accurately enough for clinical use. *Reprod Biol Endocrinol* 2017; **15**: 33 [PMID: 28449669 DOI: 10.1186/s12958-017-0251-8]
6 **Miyagi Y**, Fujiwara K, Oda T, Miyake T, Coleman RL. Development of New Method for the Prediction of Clinical Trial Results Using Compressive Sensing of Artificial Intelligence. *J Biostat Biometric App* 2018; **3**: 203
7 **VerMilyea M**, Hall JMM, Diakiw SM, Johnston A, Nguyen T, Perugini D, Miller A, Picou A, Murphy AP, Perugini M. Development of an artificial intelligence-based assessment model for prediction of embryo viability using static images captured by optical light microscopy during IVF. *Hum Reprod* 2020; **35**: 770-784 [PMID: 32240301 DOI: 10.1093/humrep/deaa013]
8 **Miyagi Y**, Habara T, Hirata R, Hayashi N. Feasibility of artificial intelligence for predicting live birth without aneuploidy from a blastocyst image. *Reprod Med Biol* 2019; **18**: 204-211 [PMID: 30996684 DOI: 10.1002/rmb2.12267]
9 **Fukushima K**. Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern* 1980; **36**: 193-202 [PMID: 7370364 DOI: 10.1007/BF00344251]
10 **Hubel DH**, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 1968; **195**: 215-243 [PMID: 4966457 DOI: 10.1113/jphysiol.1968.sp008455]
11 **Hubel DH**, Wiesel TN. Receptive fields of single neurones in the cat's striate cortex. *J Physiol* 1959; **148**: 574-591 [PMID: 14403679 DOI: 10.1113/jphysiol.1959.sp006308]
12 **Schmidhuber J**. Deep learning in neural networks: an overview. *Neural Netw* 2015; **61**: 85-117 [PMID: 25462637 DOI: 10.1016/j.neunet.2014.09.003]
13 **Miyagi Y**, Habara T, Hirata R, Hayashi N. Feasibility of predicting live birth by combining conventional embryo evaluation with artificial intelligence applied to a blastocyst image in patients classified by age. *Reprod Med Biol* 2019; **18**: 344-356 [PMID: 31607794 DOI: 10.1002/rmb2.12284]
14 **Miyagi Y**, Habara T, Hirata R, Hayashi N. Feasibility of deep learning for predicting live birth from a blastocyst image in patients classified by age. *Reprod Med Biol* 2019; **18**: 190-203 [PMID: 30996683 DOI: 10.1002/rmb2.12266]
15 **Miyagi Y**, Miyake T. Potential of deep learning for predicting fetal weight of Japanese. *Acta Med Okayama* 2020; In press
16 **Weiss RV**, Clapauch R. Female infertility of endocrine origin. *Arq Bras Endocrinol Metabol* 2014; **58**: 144-152 [PMID: 24830591 DOI: 10.1590/0004-2730000003021]
17 **Shirasuna K**, Iwata H. Effect of aging on the female reproductive function. *Contracept Reprod Med* 2017; **2**: 23 [PMID: 29201428 DOI: 10.1186/s40834-017-0050-9]
18 **Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology**. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Hum Reprod* 2011; **26**: 1270-1283 [PMID: 21502182 DOI: 10.1093/humrep/der037]
19 **Bengio Y**, Courville A, Vincent P. Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Mach Intell* 2013; **35**: 1798-1828 [PMID: 23787338 DOI: 10.1109/TPAMI.2013.50]
20 **LeCun YA**, Bottou L, Orr GB, Müller KR. Efficient Backprop. In: Montavon G, Orr GB, Müller KR, editors. Neural networks: Tricks of the trade. Berlin, Heidelberg: Springer, 2012: 9-48 [DOI: 10.1007/978-3-642-35289-8_3]
21 **LeCun Y**, Bottou L, Bengio Y, Haffner P. Gradient-Based Learning Applied to Document Recognition. *Proc IEEE* 1998; **86**: 2278-2324 [DOI: 10.1109/5.726791]
22 **LeCun Y**, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel, LD. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput* 1989; **1**: 541-551 [DOI:

10.1162/neco.1989.1.4.541]
23 **Serre T**, Wolf L, Bileschi S, Riesenhuber M, Poggio T. Robust object recognition with cortex-like mechanisms. *IEEE Trans Pattern Anal Mach Intell* 2007; **29**: 411-426 [PMID: 17224612 DOI: 10.1109/TPAMI.2007.56]
24 **Wiatowski T**, Bölcskei H. A Mathematical Theory of Deep Convolutional Neural Networks for Feature Extraction. *IEEE Trans Inf Theory* 2017; **64**: 1845-1866 [DOI: 10.1109/TIT.2017.2776228]
25 **Srivastava N**, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J Mach Learn Res* 2014; **15**: 1929-1958
26 **Nowlan SJ**, Hinton GE. Simplifying Neural Networks by Soft Weight-Sharing. Neural Comput 1992; 4: 473-493 [DOI: 10.1162/neco.1992.4.4.473]
27 **Bengio Y**. Learning Deep Architectures for AI, Founds Trends® in Mach Learn. Boston: Now Publishers Inc, 2009: 1-127
28 **Mutch J**, Lowe DG. Object Class Recognition and Localization Using Sparse Features with Limited Receptive Fields. *Int J Comput Vis* 2008; **80**: 45-57 [DOI: 10.1007/s11263-007-0118-0]
29 **Neal RM**. Connectionist Learning of Belief Networks. *Artif Intell* 1992; **56**: 71-113 [DOI: 10.1016/0004-3702(92)90065-6]
30 **Ciresan DC**, Meier U, Masci J, Gambardella LM, Schmidhuber J. Flexible, High Performance Convolutional Neural Networks for Image Classification. Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence; July 16-22; Barcelona, Spain. Menlo Park: AAAI Press, 2011: 1237-1242
31 **Scherer D**, Müller A, Behnke S. Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition. In: Diamantaras K, Duch W, Iliadis LS, editors. Artificial Neural Networks – ICANN 2010. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2010: 92-101 [DOI: 10.1007/978-3-642-15825-4_10]
32 **Huang FJ**, LeCun Y. Large-Scale Learning with Svm and Convolutional for Generic Object Categorization. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2006 June 17-22; New York, USA. IEEE, 2006: 284-291 [DOI: 10.1109/CVPR.2006.164]
33 **Jarrett K**, Kavukcuoglu K, Ranzato M, LeCun Y. What Is the Best Multi-Stage Architecture for Object Recognition? In: 2009 IEEE 12th international conference on computer vision; 2009 Sep 29-Oct 2; Kyoto, Japan; IEEE, 2009: 2146-2153 [DOI: 10.1109/ICCV.2009.5459469]
34 **Zheng Y**, Liu Q, Chen E, Ge Y, Zhao JL. Time Series Classification Using Multi-Channels Deep Convolutional Neural Networks. In: Li F, Li G, Hwang S, Yao B, Zhang Z, editors. Web-Age Information Management. WAIM 2014. Lecture Notes in Computer Science. Cham: Springer, 2014: 298-310 [DOI: 10.1007/978-3-319-08010-9_33]
35 **Mnih V**, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G, Petersen S, Beattie C, Sadik A, Antonoglou I, King H, Kumaran D, Wierstra D, Legg S, Hassabis D. Human-level control through deep reinforcement learning. *Nature* 2015; **518**: 529-533 [PMID: 25719670 DOI: 10.1038/nature14236]
36 **Szegedy C**, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going Deeper with Convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition; 2015 June 7-12; Boston, USA. Computer Vision Foundation, 2015: 1-9
37 **Glorot X**, Bordes A, Bengio Y. Deep Sparse Rectifier Neural Networks. Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS) 2011; 2011 Apr 11-13; Lauderdale, USA. AISTATS, 2011: 315-323
38 **Nair V**, Hinton GE. Rectified Linear Units Improve Restricted Boltzmann Machines. Proceedings of the 27th international conference on machine learning (ICML-10); 2010 June 21-24; Haifa, Israel. Omni press, 2010: 807-814
39 **Ioff**, S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. Available from: https://arxiv.org/abs/1502.03167v3
40 **Krizhevsky A**, Sutskever I, Hinton GE. Imagenet Classification with Deep Convolutional Neural Networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. Proceedings of the 25th International Conference on Neural Information Processing Systems; 2012 Dec 3-8; Lake Tahoe, USA. Red Hook: Curran Associates Inc., 2012: 1097-1105
41 **Bridle JS**. Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition. In: Soulié FF, Hérault J, editors. Neurocomputing. Berlin, Heidelberg: Springer, 1990: 227-236 [DOI: 10.1007/978-3-642-76153-9_28]
42 **Kohavi R**. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In: Proceedings of the 14th international joint conference on Artificial intelligence; 1995 Aug 20-25; Montreal, Canada. San Francisco: Morgan Kaufmann Publishers Inc, 1995: 1137-1145
43 **Schaffer C**. Selecting a Classification Method by Cross-Validation. *Mach Learn* 1993; **13**: 135-143 [DOI: 10.1007/BF00993106]
44 **Refaeilzadeh P**, Tang L, Liu H. Cross-validation. In: Liu, L, Özsu, MT, editors. Encyclopedia of database systems. New York: Springer, 2009: 532-538
45 **Youden WJ**. Index for rating diagnostic tests. *Cancer* 1950; **3**: 32-35 [PMID: 15405679 DOI: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3]
46 **Unal I**. Defining an Optimal Cut-Point Value in ROC Analysis: An Alternative Approach. *Comput Math Methods Med* 2017; **2017**: 3762651 [PMID: 28642804 DOI: 10.1155/2017/3762651]
47 **Litjens G**, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, Hulsbergen-van de Kaa C, Bult P, van Ginneken B, van der Laak J. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep* 2016; **6**: 26286 [PMID: 27212078 DOI: 10.1038/srep26286]
48 **Ortiz A**, Munilla J, Górriz JM, Ramírez J. Ensembles of Deep Learning Architectures for the Early Diagnosis of the Alzheimer's Disease. *Int J Neural Syst* 2016; **26**: 1650025 [PMID: 27478060 DOI: 10.1142/S0129065716500258]
49 **Gil D**, Johnsson M, Chamizo JMG, Paya AS, Fernandez DR. Application of Artificial Neural Networks in the Diagnosis of Urological Dysfunctions. *Expert Syst Appl* 2009; **36**: 5754-5760 [DOI:

10.1016/j.eswa.2008.06.065]
50 **Simões PW**, Izumi NB, Casagrande RS, Venson R, Veronezi CD, Moretti GP, da Rocha EL, Cechinel C, Ceretta LB, Comunello E, Martins PJ, Casagrande RA, Snoeyer ML, Manenti SA. Classification of images acquired with colposcopy using artificial neural networks. *Cancer Inform* 2014; **13**: 119-124 [PMID: 25374454 DOI: 10.4137/CIN.S17948]
51 **Sato M**, Horie K, Hara A, Miyamoto Y, Kurihara K, Tomio K, Yokota H. Application of deep learning to the classification of images from colposcopy. *Oncol Lett* 2018; **15**: 3518-3523 [PMID: 29456725 DOI: 10.3892/ol.2018.7762]
52 **Miyagi Y**, Takehara K, Miyake T. Application of deep learning to the classification of uterine cervical squamous epithelial lesion from colposcopy images. *Mol Clin Oncol* 2019; **11**: 583-589 [PMID: 31692958 DOI: 10.3892/mco.2019.1932]
53 **Miyagi Y**, Takehara K, Nagayasu Y, Miyake T. Application of deep learning to the classification of uterine cervical squamous epithelial lesion from colposcopy images combined with HPV types. *Oncol Lett* 2020; **19**: 1602-1610 [PMID: 31966086 DOI: 10.3892/ol.2019.11214]
54 **Olczak J**, Fahlberg N, Maki A, Razavian AS, Jilert A, Stark A, Sköldenberg O, Gordon M. Artificial intelligence for analyzing orthopedic trauma radiographs. *Acta Orthop* 2017; **88**: 581-586 [PMID: 28681679 DOI: 10.1080/17453674.2017.1344459]
55 **Khosravi P**, Kazemi E, Zhan Q, Malmsten JE, Toschi M, Zisimopoulos P, Sigaras A, Lavery S, Cooper LAD, Hickman C, Meseguer M, Rosenwaks Z, Elemento O, Zaninovic N, Hajirasouliha I. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. *NPJ Digit Med* 2019; **2**: 21 [PMID: 31304368 DOI: 10.1038/s41746-019-0096-y]
56 **Sanders B**. Uterine factors and infertility. *J Reprod Med* 2006; **51**: 169-176 [PMID: 16674011]
57 **Ikhena DE**, Bulun SE. Literature Review on the Role of Uterine Fibroids in Endometrial Function. *Reprod Sci* 2018; **25**: 635-643 [PMID: 28826369 DOI: 10.1177/1933719117725827]
58 **Taylor E**, Gomel V. The uterus and fertility. *Fertil Steril* 2008; **89**: 1-16 [PMID: 18155200 DOI: 10.1016/j.fertnstert.2007.09.069]
59 **Tomassetti C**, D'Hooghe T. Endometriosis and infertility: Insights into the causal link and management strategies. *Best Pract Res Clin Obstet Gynaecol* 2018; **51**: 25-33 [PMID: 30245115 DOI: 10.1016/j.bpobgyn.2018.06.002]
60 **Christ JP**, Gunning MN, Palla G, Eijkemans MJC, Lambalk CB, Laven JSE, Fauser BCJM. Estrogen deprivation and cardiovascular disease risk in primary ovarian insufficiency. *Fertil Steril* 2018; **109**: 594-600.e1 [PMID: 29605405 DOI: 10.1016/j.fertnstert.2017.11.035]
61 **Arronet GH**, Eduljee SY, O'Brien JR. A nine-year survey of Fallopian tube dysfunction in human infertility. *Fertil Steril* 1969; **20**: 903-918 [PMID: 5361465 DOI: 10.1016/s0015-0282(16)37205-3]
62 **Segars JH**, Herbert CM 3rd, Moore DE, Hill GA, Wentz AC, Winfield AC. Selective fallopian tube cannulation: initial experience in an infertile population. *Fertil Steril* 1990; **53**: 357-359 [PMID: 2298319 DOI: 10.1016/s0015-0282(16)53296-8]
63 **Practice Committee of the American Society for Reproductive Medicine**. The role of immunotherapy in in vitro fertilization: a guideline. *Fertil Steril* 2018; **110**: 387-400 [PMID: 30098685 DOI: 10.1016/j.fertnstert.2018.05.009]
64 **Hong YH**, Kim SJ, Moon KY, Kim SK, Jee BC, Lee WD, Kim SH. Impact of presence of antiphospholipid antibodies on *in vitro* fertilization outcome. *Obstet Gynecol Sci* 2018; **61**: 359-366 [PMID: 29780778 DOI: 10.5468/ogs.2018.61.3.359]
65 **Moreno I**, Simon C. Relevance of assessing the uterine microbiota in infertility. *Fertil Steril* 2018; **110**: 337-343 [PMID: 30098680 DOI: 10.1016/j.fertnstert.2018.04.041]
66 **Kroon SJ**, Ravel J, Huston WM. Cervicovaginal microbiota, women's health, and reproductive outcomes. *Fertil Steril* 2018; **110**: 327-336 [PMID: 30098679 DOI: 10.1016/j.fertnstert.2018.06.036]
67 **Campbell A**, Fishel S, Bowman N, Duffy S, Sedler M, Thornton S. Retrospective analysis of outcomes after IVF using an aneuploidy risk model derived from time-lapse imaging without PGS. *Reprod Biomed Online* 2013; **27**: 140-146 [PMID: 23683847 DOI: 10.1016/j.rbmo.2013.04.013]
68 **LeCun Y**, Haffner P, Bottou L, Bengio Y. Object Recognition with Gradient-Based Learning. In: Forsyth DA, Mundy JL, di Gesú V, Cipolla R. Shape, contour and grouping in computer vision. Lecture Notes in Computer Science 1681. Berlin, Heidelberg: Springer, 1999: 319-345 [DOI: 10.1007/3-540-46805-6_19]
69 **He K**, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016 June 27-30; Las Vegas, USA. IEEE, 2016: 770–778 [DOI: 10.1109/CVPR.2016.90]
70 **Hu J**, Shen L, Sun G. Squeeze-and-Excitation Networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018 June 18-23; Salt Lake City, USA. IEEE, 2018: 7132–7141 [DOI: 10.1109/CVPR.2018.00745]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com