Received: 14 February 2024 | Accepted: 27 May 2024

DOI: 10.1002/rmb2.12590

Reproductive Medicine and Biology WILEY

MINI REVIEW

Potential for artificial intelligence in medicine and its application to male infertility

Hideyuki Kobayashi

Department of Urology, Toho University School of Medicine, Tokyo, Japan

Correspondence
Hideyuki Kobayashi, Department of Urology, Toho University School of Medicine, 6-11-1, Omori-Nishi, Ota-ku, Tokyo 143-8541, Japan.
Email: hideyukk@med.toho-u.ac.jp

Funding information
Japan Society for the Promotion of Science, Grant/Award Number: JP22K09486

Abstract

Background: The third AI boom, which began in 2010, has been characterized by the rapid evolution and diversification of AI and marked by the development of key technologies such as machine learning and deep learning. AI is revolutionizing the medical field, enhancing diagnostic accuracy, surgical outcomes, and drug production.

Methods: This review includes explanations of digital transformation (DX), the history of AI, the difference between machine learning and deep learning, recent AI topics, medical AI, and AI research in male infertility.

Main Findings (Results): In research on male infertility, I established an AI-based prediction model for Johnsen scores and an AI predictive model for sperm retrieval in non-obstructive azoospermia, both by no-code AI.

Conclusions: AI is making constant progress. It would be ideal for physicians to acquire a knowledge of AI and even create AI models. No-code AI tools have revolutionized model creation, allowing individuals to independently handle data preparation and model development. Previously a team effort, this shift empowers users to craft customized AI models solo, offering greater flexibility and control in the model creation process.

KEYWORDS
artificial intelligence, deep learning, Johnsen score, machine learning, micro-TESE

1 | INTRODUCTION

The third AI boom, which began in 2010, has been characterized by its rapid evolution and diversification and is marked by the development of key technologies such as machine learning, deep learning, and genetic algorithms.[1] Particularly in industry, AI is moving into an implementation-oriented phase. With regard to medical AI, solutions are being developed with implementation in mind, and many startup companies have been founded. Since the number of physicians is expected to decrease in developed countries as the problem of declining birthrates becomes more pronounced, I believe it is necessary to transform medical services through the widespread use of digital transformation (DX).

I will outline ways AI is being applied to health care. The first is image recognition using deep learning. Image recognition includes image classification, object detection, and segmentation. Image classification is a technique used to identify and classify what is in an image. Object detection is a technique that identifies the location of an object in an image and what it is. Segmentation is a technique used to identify the boundaries of objects in pixel units. The second is classification based on machine learning. Machine learning classification requires the creation of variables called features. In achieving accuracy, creating features is the most difficult process. The third is regression analysis, which is used to obtain a prediction equation for the numerical value that serves as the objective variable when dealing with numerical data. The fourth is text analysis, which includes sentence analysis and spoken language treatment.

This review article explains what AI is, how medical AI research is being conducted, and examples of AI applications to male infertility.

2 | WHAT IS DX?

Driven by emerging technologies, DX is a strategic shift in organizations that fundamentally changes how value is delivered to customers.[2] It involves a holistic approach to implementing new methods and technologies to enhance organizational performance and competitiveness.[3] The integration of digital technology into all aspects of a business has significant impacts on business models, operational processes, and customer experience.[4] In the context of sweeping social, economic, and technological changes, DX is a necessary and transformative process for institutions, particularly in higher education.[5]

There are three key points in promoting DX: first, develop high-level Information Technology (IT) personnel who can promote it; second, upgrade digital tools through their application on a daily basis; and third, continuously examine existing business processes and digital technologies. In promoting DX in the world of medicine, the use of machine learning, AI, and data to drive change would be important.

3 | BRIEF HISTORY OF AI

The Dartmouth Conference of 1956 is widely regarded as the birthplace of AI.[6] This event brought together key figures in AI, including John McCarthy, Marvin Minsky, Allen Newell, and Herbert Simon, who later went on to make significant contributions to the field.[7] The conference marked the beginning of a period of great optimism about the potential of AI, with the belief that human intelligence could be replicated in machines.[8] However, there have been significant challenges in achieving this goal, with the complexity of human behavior proving to be a major obstacle.[9] Despite these challenges, significant progress has been made in the field of AI in the past 50 years, and a wide range of tools and techniques has emerged from research.[9]

The first AI boom is often associated with the 1956 Dartmouth Conference, where the term "artificial intelligence" was coined and the field of AI was formally established.[10,11] The ability to provide solutions to specific problems led to the first AI boom, with "reasoning" and "search" becoming the major keywords during it. Currently, we are considered to be in the third AI boom. Beginning in 2010, it has been marked by rapid evolution and diversification, particularly in medical diagnosis and diagnostic imaging.[12] In the medical field, the strength of deep learning is that various types of image information, such as CT and MRI images, can be learned and automatically described features. A major advantage of this is that AI can notice lesions that would otherwise be overlooked by physicians. In addition, the potential economic impact of the third boom has also been highlighted and it has been suggested that it could drive the development of "apps" and contribute to a new economic boom.[13]

4 | MACHINE LEARNING AND DEEP LEARNING

Machine learning, a key component of AI, involves training computers to perform tasks using data and algorithms.[14] It is a scientific discipline focused on developing algorithms that enable computers to evolve behaviors based on empirical data.[15] Machine learning has diverse applications, from character recognition to medical diagnosis and product classification.[16] It allows computers to learn, grow, and develop independently when presented with new data and has been applied in various domains such as self-driving cars, online recommendation engines, and cyber fraud detection.[17]

Deep learning, a form of machine learning, is a powerful technique for constructing high-dimensional predictors in input–output models.[18] It operates through learning in multiple levels of abstraction, making it particularly effective for pattern learning and recognition.[19] The algorithms used, also known as neural networks, are inspired by the structure of the brain and have been successful in complex problem-solving, such as in computer vision and language modeling.[20] Deep learning, a key component of AI, has seen significant advances in recent years, particularly in the development of deep architectures and learning algorithms.[21]

Thus, deep learning is included in the machine learning framework, where AI is defined as a larger framework that includes machine learning and deep learning.

5 | RECENT ADVANCES IN AI

5.1 | Foundation models

Foundation models, such as BERT, DALL-E, and GPT-3, are large language models that have revolutionized various fields of research.[22,23] BERT, a powerful language representation model, has been shown to significantly improve conversation modeling,[24] language understanding,[25] and entity ranking.[26] DALL-E is a powerful image generation model developed by OpenAI, an AI research and deployment company, which is capable of creating diverse and realistic images from textual descriptions. It uses a combination of deep learning and natural language processing to understand and interpret the input text, and then generate corresponding images. This model has been trained on a wide range of images and text, allowing it to produce a variety of visual concepts and designs. DALL-E has the potential to revolutionize the field of computer-generated imagery and has already demonstrated its capabilities in creating unique and imaginative visuals. Also developed by OpenAI, GPT-3 is a powerful autoregressive language model with 175 billion parameters, capable of performing a wide range of natural language processing (NLP) tasks.

Chat GPT is a powerful language model based on generative pretrained transformer (GPT) architecture and is capable of generating human-like text.[27] Having the ability to detect and correct grammatical errors, improve text coherence and clarity, and generate additional content,[28] it has shown promise in academic editing. In the field of medical education, Chat GPT has been used to generate virtual patient simulations, quizzes, and curriculums, highlighting its potential in this area.[29] However, it is important to note that the model has limitations and challenges that need to be addressed.

5.2 | Prediction model for 3D structure of proteins

AlphaFold2 is a groundbreaking protein structure prediction tool that has revolutionized the field of molecular biology, developed by DeepMind, a subsidiary of Alphabet Inc. It has been shown to be highly accurate, with a median domain GDT_TS of 92.4, and has significantly improved the state-of-the-art in protein structure prediction.[30] GDT_TS (Global Distance Test Total Score) is a measure used to assess the accuracy of protein structure predictions. It has been shown to be highly accurate, with a median domain GDT_TS of 92.4, meaning that the predicted structure is 92.4% similar to the true structure. The use of this tool has been extended to generate ensembles of structures, providing a more comprehensive view of protein dynamics.[31] Furthermore, AlphaFold2 has been used to predict the appearance of functional adaptations in evolution, demonstrating its versatility and potential impact on evolutionary biology.[32]

Thus, AI is constantly evolving, and a major strength is that it is very effective when there is a large amount of data for a single, uniform task.

5.3 | What is medical AI?

AI is revolutionizing the medical field, enhancing diagnostic accuracy, surgical outcomes, and drug production.[33] It is particularly impactful in medical image analysis, reducing the gap between research and deployment. AI's potential to improve health care is vast, but it also presents challenges such as data scarcity and racial bias.[34] Despite these challenges, AI is expected to transform primary care and have potential applications in medical imaging analysis, drug discovery, and personalized medicine.[33]

However, its use in health care is still in the early stages, and it may take until the 2030s for it to fully revolutionize the industry.[35] Reasons why the transformation of health care through AI will not happen immediately include the following: the data necessary for AI has not been digitized, technology and systems to solve the current situation where doctors are busy with administrative work have not caught up, and AI built based on data from a single facility is not accurate enough for analysis of data from other facilities.

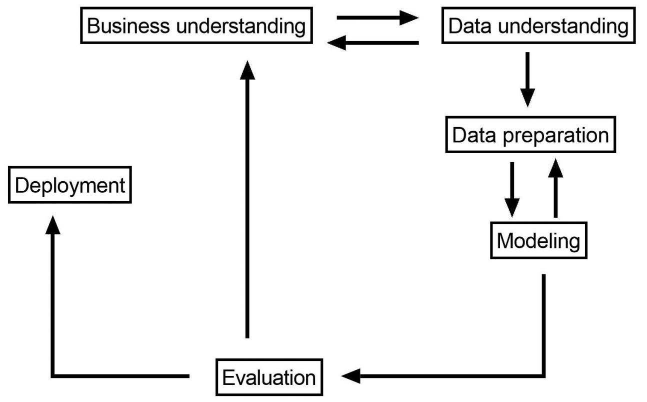

FIGURE 1 Six phases of CRISP-DM.

5.4 | Developing medical AI models

The key to successfully developing medical AI models is having AI engineers work closely together with users. In addition, physicians should attempt AI programming as much as possible. Furthermore, development in accordance with the Cross-Industry Standard Process for Data Mining (CRISP-DM) will allow projects to proceed smoothly. CRISP-DM is a widely used process model for data mining projects. It was developed by a consortium of industry experts and is designed to be industry and technology neutral.[36] The CRISP-DM model consists of six phases (Figure 1): 1. business understanding, 2. data understanding, 3. data preparation, 4. modeling, 5. evaluation, and 6. deployment.

5.4.1 | Business understanding

This is the phase in which the AI developer works with the user and the marketing manager to understand the business situation. In this case, the user is assumed to be a physician. Close communication between the AI developer and the user is critical.

5.4.2 | Data understanding

The AI developer has to be able to understand the data. The user, the physician, needs to teach the AI developer the respective terminology.

5.4.3 | Data preparation

The AI developer converts the data provided by the user into a format suitable for AI development. This operation is called data cleansing.

5.4.4 | Modeling

The AI developer creates an AI model that meets the customer's requirements, taking into account data items, quantity, and quality.

5.4.5 | Evaluation

The developed AI model is evaluated. If the value of the evaluation index does not reach the target, improve the model or consider a different method of proceeding.

5.4.6 | Deployment

Once evaluation of the model is complete and goals are met, deploy to the field. Consider the User Interface (UI)/User Experience (UX) layout so that the tool is easy to use.

The CRISP-DM model also includes a hierarchical and iterative approach, allowing for flexibility in its application.[37] For example, It has been extended to address specific challenges in various domains, such as medicine[38] and machine learning.[39] These extensions provide a more detailed framework for conducting data mining analysis in these specialized fields.

6 | AI MOVING FROM "CREATION" TO "USE"

6.1 | No-code AI

No-code AI is a service that enables machine learning without the need for coding, or the help of machine learning engineers, data scientists, or other specialists. It is also called automated machine learning (AutoML). The emergence of no-code AI tools has been transforming various domains, including industrial process engineering,[40] business process automation,[41] and software development.[42,43]

No-code AI allows physicians with no expertise in AI to create AI models. As an example, in a study that comprehensively analyzed the performance and feature set of six no-code AI platforms using four representative medical imaging datasets to create image classification models, those of Amazon and Google demonstrated higher classification performance, with Amazon's performing significantly better than Apple's.

Although the study did not show significant performance differences among the leading platforms, they differed significantly in terms of the critically important evaluation features available. The authors believed that code-free deep learning platforms have the potential to improve access to deep learning for both clinicians and biomedical researchers.[44]

Until now, creating an AI model has required a team and multiple people to work on it. With no-code; however, a single person can be responsible for preparing data and creating AI models. This gives the user the advantage of being able to create the AI model he or she wants to create by him/herself. In addition, I think that no-code could automate the modeling in CRISP-DM.

6.2 | Research on application of AI to male infertility

We have created an image recognition AI model and a predictive analytics AI model for male infertility, which are described in the following.

6.3 | AI-based prediction model for Johnsen scores

Patients with azoospermia need to undergo testicular sperm extraction (TESE) to obtain mature sperm. This comprises conventional TESE (for obstructive type) and microdissection TESE (micro-TESE) (for non-obstructive type). In addition, the condition of the testis is checked by collecting a piece of testis tissue in TESE, where the Johnsen score is an effective means of evaluating histological features of the testis.

Evaluation of testicular histology is an important end point in understanding male infertility. Spermatogenic potential is evaluated by the Johnsen score. Published in 1970, it is still in use today.[45] The score ranges from 1 to 10 points; the higher the score, the greater the spermatogenic potential, with spermatozoa observable for a score of 8 or higher. Johnsen scores are determined through examination by pathologists. However, the time taken for patients to know the result not only places an emotional burden on them but also hinders treatment plans.

Therefore, we created a computer vision algorithm for classifying Johnsen scores using a no-code AI approach.[46] We defined four labels: Johnsen score 1–3, 4–5, 6–7, and 8–10, to distinguish Johnsen score ranges. Briefly, we obtained a dataset of 7155 images of testicular tissue at magnification 400X and developed an AI-based algorithm for evaluating Johnsen scores, with an AUC of 82.6% (Figure 2A). We also obtained a dataset of 9822 expansion images for 5.0 X 5.0 cm cutouts and developed an AI-based algorithm for evaluating Johnsen scores, with an AUC of 99.5% (Figure 2B). The algorithms have the potential to support pathologists' evaluations and could be used as an alternative to traditional Johnsen scoring.

6.4 | AI predictive model for sperm retrieval in non-obstructive azoospermia

The sperm extraction rate by micro-TESE in non-obstructive azoospermia (NOA) was reported to be about 34.0% in Japan, which is not very high.[47] In addition, there had been no factors predictive of sperm retrieval. Therefore, we attempted to create an AI model for predicting sperm retrieval in patients with NOA before undergoing micro-TESE. Briefly, we performed sperm retrieval in patients with NOA and developed an AI model using Prediction One, which does not require coding, to make an AI prediction model for predicting

FIGURE 2 AI-based algorithm for predicting Johnsen scores.

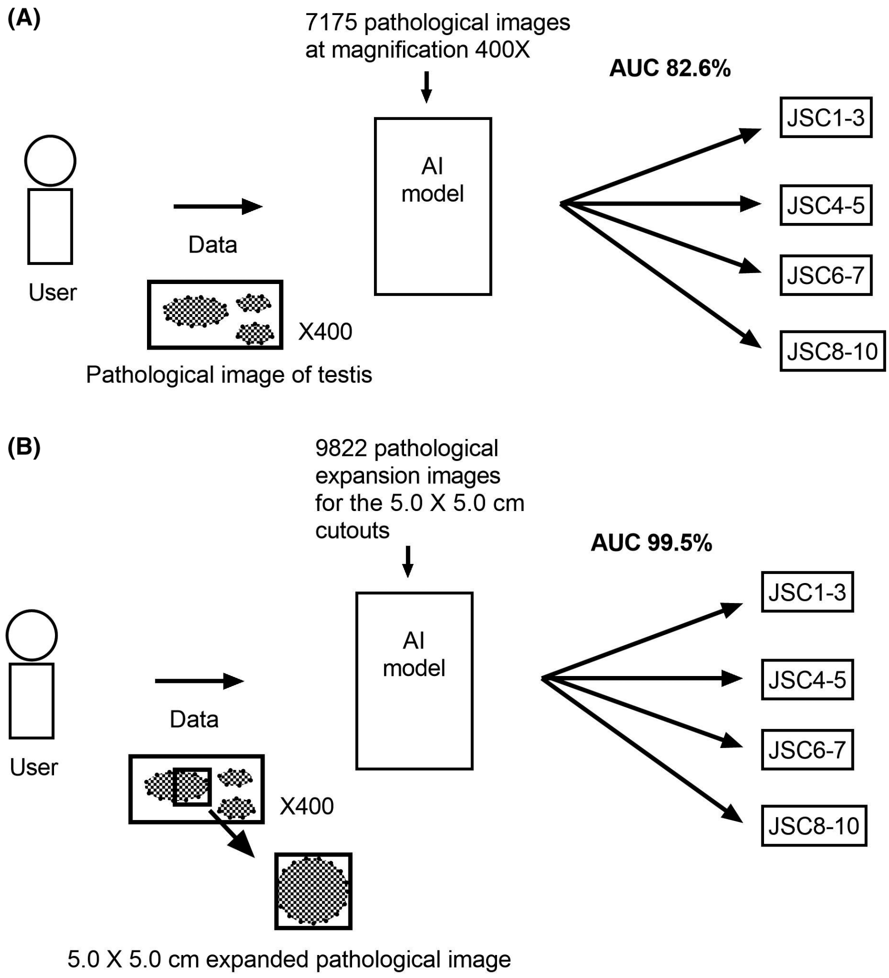

FIGURE 3 AI model for predicting sperm retrieval in patients with NOA before undergoing micro-TESE.

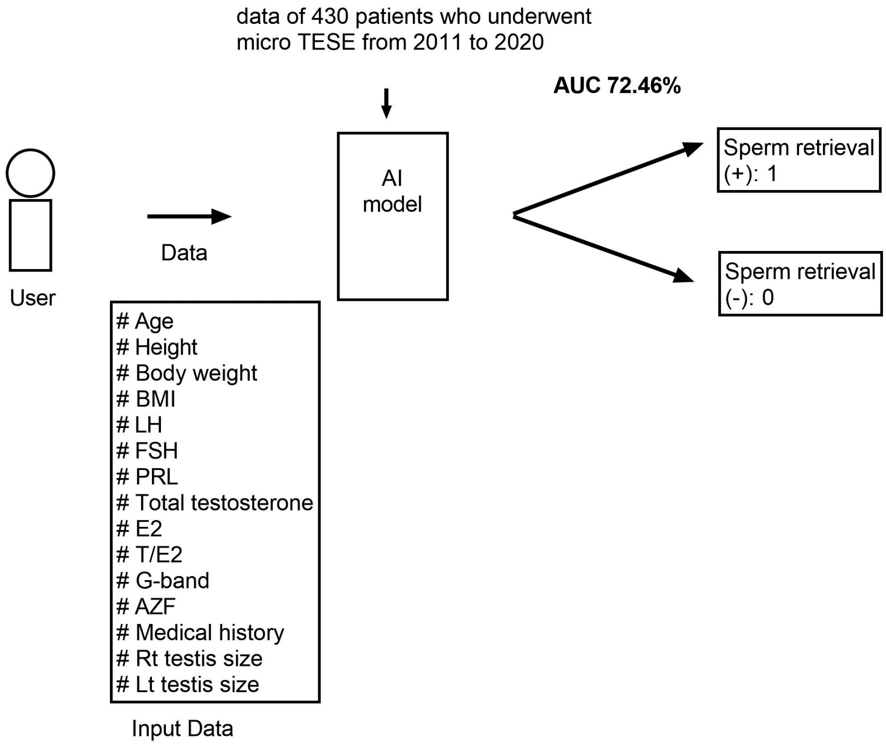

TABLE 1 AI models in male infertility.

Author	Target	Title
Tsai et al. (2020)[51]	Smartphone application that measures sperm motility	Web- and artificial intelligence-based image recognition for sperm motility analysis
Hicks et al. (2019)[52]	Analysis of sperm video	Machine Learning-based Analysis of Sperm Videos and Participant Data for Male Fertility Prediction
Riegler et al. (2021)[53]	Analysis of sperm video	Artificial intelligence in the fertility clinic: status, pitfalls and possibilities
Chandra et al. (2022)[54]	Analysis of sperm morphology	Prolificacy Assessment of Spermatozoan via State-of-the Art Deep Learning Frameworks
Ilhan et al. (2022)[55]	Analysis of sperm morphology	Sperm morphology analysis by using the fusion of two-stage fine-tuned deep networks
Ilhan, et al. (2020)[56]	Analysis of sperm morphology	A fully automated hybrid human sperm detection and classification system based on mobile-net and the performance comparison with conventional methods
Movahed et al. (2019)[57]	Analysis of sperm morphology	Automatic segmentation of sperm parts in microscopic images of human semen smears using concatenated learning approaches
Abbasi et al. (2021)[58]	Selection of optimum sperm for ICSI	Effect of deep transfer and multi-task learning on sperm abnormality detection
Javadi et al. (2019)[59]	Selection of optimum sperm for ICSI	A novel deep learning method for automatic assessment of human sperm images
Zeadna et al. (2020)[60]	Prediction of obtaining spermatozoa from testicular biopsies	Prediction of sperm extraction in non-obstructive azoospermia patients: a machine-learning perspective

Abbreviation: ICSI, intracytoplasmic sperm injection.

the possibility of sperm retrieval before micro-TESE. To do this, we obtained data from the medical records of 430 patients who underwent micro-TESE from 2011 to 2020. The parameters extracted were age, height, body weight, body mass index, LH, FSH, PRL, total testosterone, E2, T/E2, sperm retrieval, G-band, AZF, medical history, Rt testis, and Lt testis. The AI model achieved an acceptable AUC of 72.46%. T/E2 ratios were found to be the most important variable for predicting sperm retrieval[48] (Figure 3).

6.5 | Application of AI in reproductive medicine

Recent reviews state that AI algorithms have become ubiquitous in the field of reproductive medicine (Table 1).[49,50]

Below, I list some examples of the numerous published AI models of male infertility.

Sperm analysis

1. Smartphone application for measuring sperm motility[51]
2. Analysis of sperm videos[52,53]
3. Analysis of sperm morphology[54–57]
4. Selection of optimum sperm for intracytoplasmic sperm injection (ICSI)[58,59]
5. Prediction of spermatozoa retrieval from testicular biopsies[60]

It can be seen that most AI research in male infertility has been preceded by analyses on sperm morphology using smartphone apps or video and sperm selection in ICSI. Our group was the first to report an automatic AI discrimination model for the Johnsen score, which determines sperm maturity in testicular pathological tissue.[46] In addition, although a sperm retrieval AI prediction model for conventional TESE in patients with NOA had been previously reported,[60] our study is the first on a sperm retrieval AI prediction model for micro-TESE in NOA patients.[48]

7 | CONCLUSION

Medical AI research should be led by medical professionals. Although AI engineers can build AI models for image classification, which is not a difficult task, they cannot deduce specific needs. Only physicians can make comprehensive judgments for determining specific diseases, for which AI will be used and satisfying on-site needs such as improving operational efficiency. Thus, it would be ideal for physicians themselves to acquire knowledge of AI and create AI models. Finally, no-code AI tools enable individuals to independently prepare data and develop AI models, a task previously requiring a team. This empowers users to create customized AI models without external assistance, streamlining the process and increasing accessibility.

ACKNOWLEDGMENTS

Hideyuki Kobayashi has received support in the form of a Grant-in-Aid for Scientific Research (C) from the Japan Society for the Promotion of Science (JSPS) (JSPS KAKENHI Grant Number JP22K09486). We are particularly grateful to Alexander Cox for his painstaking work as medical editor.

CONFLICT OF INTEREST STATEMENT

The author declares no conflict of interest.

HUMAN/ANIMAL RIGHTS

This article does not contain any studies with human or animal subjects performed by the author.

ORCID

Hideyuki Kobayashi https://orcid.org/0000-0003-1381-1665

REFERENCES

1. Miyazaki K, Sato R, editors. Analyses of the technological accumulation over the 2nd and the 3rd AI boom and the issues related to AI adoption by firms. 2018 Portland International Conference on Management of Engineering and Technology (PICMET) 2018: 1–7.
2. Tang D. What is digital transformation? EDPACS. 2021;64(1):9–13.
3. Pihir I, Tomičić-Pupek K, Furjan MT, editors. Digital transformation insights and trends. 2018.
4. Gebayew C, Hardini IR, Panjaitan GHA, Kurniawan NB, Systems SJICoIT, Innovation. A systematic literature review on digital transformation. 2018; 260–5.
5. Abd-Rabo A, Hashaikeh S. The digital transformation revolution. Int J Human Educat Res. 2021;3:124–8.
6. Klassner F. In: Lorrie Faith C, editor. Special issue on artificial intelligence. New York: Association for Computing Machinery; 1996.
7. Bobrow DG, Brady MJ. Artificial intelligence 40 years later. Artif Intell. 1998;103:1–4.
8. Crevier D. AI: the tumultuous history of the search for artificial intelligence. New York: Basic Books, Inc; 1993.
9. Hopgood AA. The state of artificial intelligence. Advances in computers. vol. 65. Amsterdam, Netherlands: Elsevier; 2005. p. 1–75.
10. Bruderer HE, editor. The birth of artificial intelligence: first conference on artificial intelligence in Paris in 1951? HC. 2016.
11. Strickland EKJ. The turbulent past and uncertain future of AI: is there a way out of AI's boom-and-bust cycle? IEEE Spectrum. 2021;58:26–31.
12. Fujita HJ. AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. Radiol Phys Technol. 2020;13:6–19.
13. Szu HJ. The 3rd wave of AI can help develop apps generating the 4th economic booming. MOJ Appl Bio Biomech. 2019;3:49–56.
14. Orkun Baloglu SQL, Nazha A. What is machine learning? J Arch Dis Childhood. 2020;107:386–8.
15. Alpaydin E, editor. Introduction to machine learning. Advanced topics in artificial intelligence. 1992;163(5):732–34.
16. Mello RF, Ponti MA. A brief review on Machine Learning. 2018.
17. Argade D, Pawar S, Thitme V, Shelkar A. Machine learning: review. Int J Adv Res Sci Commun Technol. 2021;7(2):251–6.
18. Polson NG, Sokolov VO. Deep learning. 2018; abs/1807.07987.
19. Chen CLP. Informatics. Deep learning for pattern learning and recognition. In Proceedings of the 10th IEEE Jubilee International Symposium on Applied Computational Intelligence & Informatics 2015: 17.
20. Bisong E. *Building machine learning and deep learning models on google cloud platform. what is deep learning?* 2019.
21. Mo D, editor. *A survey on deep learning*: *one small step toward AI*. 2012.
22. Schneider J. Foundation models in brief: A historical, sociotechnical focus. 2022; abs/2212.08967.
23. Bommasani R, Hudson DA, Adeli E, Altman R, Arora S, Arx SV, et al. On the opportunities and risks of foundation models. 2021; abs/2108.07258.
24. Zhao X, Zhang Y, Guo W, Yuan X. BERT for open-domain conversation modeling. 2019:1532-6.
25. Devlin J, Chang M-W, Lee K, Toutanova K, editors. BERT: pre-training of deep bidirectional transformers for language understanding. Minneapolis, Minnesota: North American Chapter of the Association for Computational Linguistics; 2019.
26. Chatterjee S, Dietz L. Retrieval DiI. BERT-ER: Query-specific BERT entity representations for entity ranking. 2022.
27. Sarode SV, Bhamare VK. Technology E. Chat GPT and its capabilities. 2023.
28. Castillo-González W, Lepez CO, Bonardi MC. Chat GPT: a promising tool for academic editing. 2022.
29. Eysenbach GJ. The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. JMIR Med Educ. 2023;9:e46885.
30. Jumper JM, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Applying and improving AlphaFold at CASP14. Proteins. 2021;89:1711–21.
31. Vani BP, Aranganathan A, Wang D, Tiwary PJ. AlphaFold2-RAVE: From Sequence to Boltzmann Ranking. J Chem Theory Comput. 2023;19:4351–4.
32. Ponlachantra K, Suginta W, Robinson RC, Kitaoku YJB. AlphaFold2: A versatile tool to predict the appearance of functional adaptations in evolution. Bioessays. 2022;45:e2200119.
33. Liu P, Lu L, Zhang J, Huo T, Liu S, Ye ZJCMS. Application of artificial intelligence in medicine: an overview. Curr Med Sci. 2021;41:1105–15.
34. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med. 2022;28:31–8.
35. Komal, Sethi GK, Ahmad N, Rehman MB, Ibrahim Dafallaa HME, Rashid M, et al. Use of artificial intelligence in healthcare systems: state-of-the-art survey. 2021 243–8.
36. Chapman P, editor. CRISP-DM 1.0: Step-by-step data mining guide. 2000.
37. Wirth R, Hipp J. Crisp-dm: towards a standard process modell for data mining. 2000.
38. Niaksu O. CRISP data mining methodology extension for medical domain. 2015.
39. Martínez-Plumed F, Ochando LC, Ferri C, Flach PA, Hernández-Orallo J, Kull M, et al. CASP-DM: context aware standard process for data mining. 2017; abs/1709.09003.
40. Ogundare O, Araya GQ, Qamsane Y. No code AI: automatic generation of function block diagrams from documentation and associated heuristic for context-aware ML algorithm training. 2022:191–5.
41. Desmond M, Duesterwald E, Isahagian V, Muthusamy VJA. A no-code low-code paradigm for authoring business automations using natural language. 2022;abs/2207.10648.
42. Rao N, Tsay J, Kate K, Hellendoorn VJ, Hirzel MJA. AI for low-code for AI. 2023; abs/2305.20015.
43. Cabot J, Clarisó R, Menzies T. Low code for smart software development. IEEE Software. 2023;40:89–93.
44. Korot E, Guan Z, Ferraz DA, Wagner SK, Zhang G, Liu X, et al. Code-free deep learning for multi-modality medical image classification. Nat Mach Intell. 2021;3:288–98.
45. Johnsen SG. Testicular biopsy score count – a method for registration of spermatogenesis in human testes: normal values and results in 335 hypogonadal males. Hormones. 1970;1(1):2–25.
46. Ito Y, Unagami M, Yamabe F, Mitsui Y, Nakajima K, Nagao K, et al. A method for utilizing automated machine learning for histopathological classification of testis based on Johnsen scores. Sci Rep. 2021;11:9962.
47. Yumura Y, Tsujimura A, Imamoto T, Umemoto Y, Kobayashi H, Shiraishi K, et al. Nationwide survey of urological specialists regarding male infertility: results from a 2015 questionnaire in Japan. Reprod Med Biol. 2018;17(1):44–51.

48. Kobayashi H, Uetani M, Yamabe F, Mitsui Y, Nakajima K, Nagao K. AI model developed using machine learning for predicting sperm retrieval in micro-TESE for nonobstructive azoospermia patients. Andrologia. 2023;2023:1–10.
49. Gül M, Russo GI, Kandil H, Boitrelle F, Saleh R, Chung E, et al. Male infertility: new developments, current challenges, and future directions. World J Men's Health. 2024;42(3):502–517.
50. Güell E. Criteria for implementing artificial intelligence systems in reproductive medicine. Clin Exp Reprod Med. 2024;51(1):1–12.
51. Tsai VF, Zhuang B, Pong YH, Hsieh JT, Chang HC. Web- and artificial intelligence-based image recognition for sperm motility analysis: verification study. JMIR Med Inform. 2020;8(11):e20031.
52. Hicks SA, Andersen JM, Witczak O, Thambawita V, Halvorsen P, Hammer HL, et al. Machine learning-based analysis of sperm videos and participant data for male fertility prediction. Sci Rep. 2019;9(1):16770.
53. Riegler MA, Stensen MH, Witczak O, Andersen JM, Hicks SA, Hammer HL, et al. Artificial intelligence in the fertility clinic: status, pitfalls and possibilities. Hum Reprod. 2021;36(9):2429–42.
54. Chandra S, Gourisaria MK, Gm H, Konar D, Gao X, Wang T, et al. Prolificacy assessment of spermatozoan via state-of-the-art deep learning frameworks. IEEE Access. 2022;10:13715–27.
55. Ilhan HO, Serbes G. Sperm morphology analysis by using the fusion of two-stage fine-tuned deep networks. Biomed Signal Process Control. 2022;71:103246.
56. Ilhan HO, Sigirci IO, Serbes G, Aydin N. A fully automated hybrid human sperm detection and classification system based on mobile-net and the performance comparison with conventional methods. Med Biol Eng Comput. 2020;58(5):1047–68.
57. Movahed RA, Mohammadi E, Orooji M. Automatic segmentation of Sperm's parts in microscopic images of human semen smears using concatenated learning approaches. Comput Biol Med. 2019;109:242–53.
58. Abbasi A, Miahi E, Mirroshandel SA. Effect of deep transfer and multi-task learning on sperm abnormality detection. Comput Biol Med. 2021;128:104121.
59. Javadi S, Mirroshandel SA. A novel deep learning method for automatic assessment of human sperm images. Comput Biol Med. 2019;109:182–94.
60. Zeadna A, Khateeb N, Rokach L, Lior Y, Har-Vardi I, Harlev A, et al. Prediction of sperm extraction in non-obstructive azoospermia patients: a machine-learning perspective. Hum Reprod. 2020;35(7):1505–14.

How to cite this article: Kobayashi H. Potential for artificial intelligence in medicine and its application to male infertility. Reprod Med Biol. 2024;23:e12590. https://doi.org/10.1002/rmb2.12590