COMMENTARY

Does artificial intelligence have a role in the IVF clinic?

Darren J X Chow[1,2,3], **Philip Wijesinghe**[4], **Kishan Dholakia**[3,4,5,6] and **Kylie R Dunning**[1,2,3]

[1]Robinson Research Institute, School of Biomedicine, The University of Adelaide, Adelaide, Australia
[2]Australian Research Council Centre of Excellence for Nanoscale Biophotonics, The University of Adelaide, Adelaide, Australia
[3]Institute for Photonics and Advanced Sensing, The University of Adelaide, Adelaide, Australia
[4]SUPA, School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, Fife, United Kingdom
[5]School of Biological Sciences, The University of Adelaide, Adelaide, Australia
[6]Department of Physics, College of Science, Yonsei University, Seoul, South Korea

Correspondence should be addressed to K R Dunning: kylie.dunning@adelaide.edu.au

Lay summary

The success of IVF has remained stagnant for a decade. The focus of a great deal of research is to improve on the current ~30% success rate of IVF. Artificial intelligence (AI), or machines that mimic human intelligence, has been gaining traction for its potential to improve outcomes in medicine, such as cancer diagnosis from medical images. In this commentary, we discuss whether AI has the potential to improve fertility outcomes in the IVF clinic. Based on existing research, we examine the potential of adopting AI within multiple facets of an IVF cycle, including egg/sperm and embryo selection, as well as formulation of an IVF treatment regimen. We discuss both the potential benefits and concerns of the patient and clinician in adopting AI in the clinic. We outline hurdles that need to be overcome prior to implementation. We conclude that AI has an important future in improving IVF success.

Reproduction and Fertility (2021) **2** C29–C34

Introduction

The birth of the first baby conceived using *in vitro* fertilisation (IVF) occurred more than four decades ago. Since then, there have been many advances in IVF including personalised ovarian stimulation, extended embryo culture at physiological oxygen, and a move to single embryo transfer. However, for many infertile patients, the journey to parenthood via IVF is a lengthy and emotionally and financially challenging process.

For some patients, IVF will not lead to the birth of a baby. The focus of a great deal of research is to improve on the current ~30% success rate of IVF (De Geyter *et al.* 2020). Further advancements are likely in areas including patient-specific treatment regimen, improved gamete, and embryo selection, as well as improving endometrial receptivity and management of early pregnancy. Interestingly, variation in IVF success exists between operators and clinics. Such variation likely exists due to the subjectivity of current assessments. One such assessment is the grading and selection of embryos for transfer based on morphological features.

Recently, the use of artificial intelligence (AI) has gained traction in its ability to predict clinical outcomes using routinely obtained information, such as patient attributes, medical images, and blood test results. In turn, this has led to its demonstrated use in accurately classifying the severity of cancer using medical images (Esteva *et al.* 2017, Yamashita *et al.* 2021). Here, we consider whether AI has a place in the IVF clinic. Can AI deliver assessments that are objective and accurate? Will this lead to improved IVF success?

https://raf.bioscientifica.com
https://doi.org/10.1530/RAF-21-0043

Artificial intelligence; machine learning; deep learning; neural networks: what's the difference?

To explore the potential application of AI, we have to first define it, and how it differs from machine learning and deep learning technologies (Fig. 1). AI is an umbrella term that describes machines that aim to mimic human or animal cognitive capacity. Machine learning is a subset of the broader AI technology that learns its processing from data without explicit programming. It can predict outcomes for tasks that comprise many parameters and that would be too time-consuming (or impossible) for a person to perform. Deep learning is a further subset of machine learning. It utilises artificial neural networks (ANN) which mimic the architecture of neurons in the brain. Deep neural networks, those that comprise many layers of neurons, can perform any conceivable computation if the precise neuron connections are activated via training. This capacity alone has integrated deep learning across broad advances in science and technology. Deep learning can be further classified as either supervised or unsupervised. Supervised learning guides networks to generate outputs consistent with some label or ground truth, whilst unsupervised learning infers features from some inherent structure in the data. There are many ANN architectures, with increasingly more sophisticated systems being developed, from fully connected networks that can make predictions from unstructured data to convolutional networks that excel in image classification, to recurrent networks that make predictions based on temporally evolving data. An exposition on ANNs is beyond the scope of the current commentary and we direct the reader to a recent in-depth review (Fernandez *et al.* 2020). For the purpose of this article, we will use AI as the umbrella term to describe the mathematical algorithms that endeavour to automate any decision or analysis performed by a clinician or embryologist.

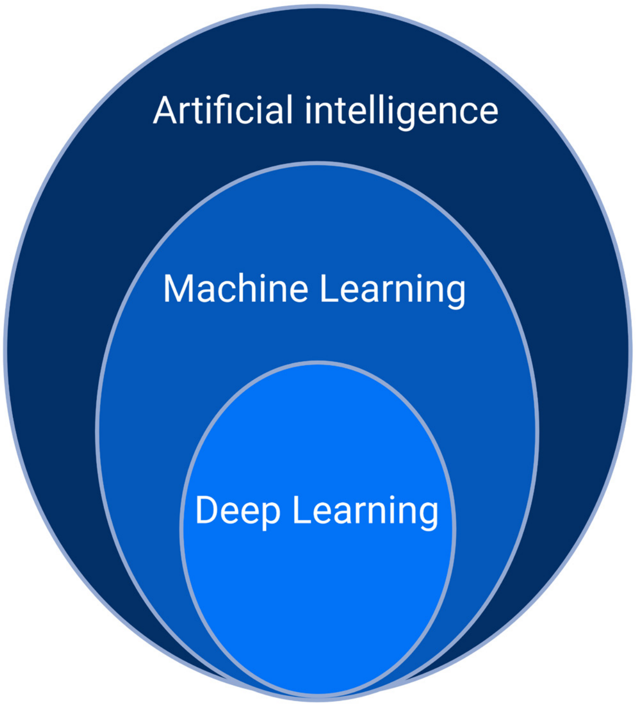

Figure 1 Hierachy of information processing theories in the computer science field. Artificial intelligence (AI) is an umbrella term attributed to the primary objective within the field of computer science to develop machines with intelligence. Machine learning describes approaches to achieve AI that learns from experience without explicit programming. Deep learning is a form of machine learning that utilises artificial neural networks to extract, process, and predict information by learning from examples. It is commonly applied in the scientific field, in particular, in image classification.

Use of AI in image classification

Optical imaging is at the fore with regard to many clinical diagnoses and is an area where AI can assist. A positive attribute for AI is that once it is trained, the inference (i.e. the processing of data through a trained network) is non-iterative and very fast to compute, obviating the need for a parameter search. This gives advantages in speed and implementation. Evaluation of networks has been well-integrated with contemporary devices and can be done on servers, using cloud services, and even on mobile phones.

When using optical imaging, clinicians are faced with the challenge of having different imaging equipment both within and between institutions. This yields images of varying quality (i.e. resolution, contrast, etc.) making standardisation of assessments challenging. For a recent review and future trends of the area of deep learning for microscopy, see Wijesinghe & Dholakia 2021. For instance, subjectivity in clinical assessment is a major issue for cancer treatment. Different operators often yield different diagnoses from the same set of images. Multi-dimensional imaging, such as time-lapse or 3D imaging, presents a further issue where the clinician is unable to analyse and interpret the data at once from one static image and must rely on memory and contextual information. Adopting AI could aid clinicians in these aspects – is the same true for clinical IVF?

Adopting AI within the fertility clinic could lead to a major increase in IVF success. For example, AI could potentially aid the embryologist in providing rapid, objective, and accurate assessment of gamete and embryo

https://raf.bioscientifica.com
https://doi.org/10.1530/RAF-21-0043

health. AI may also aid physicians in formulating an optimal, personalised fertility treatment plan based on patient characteristics. Supervised AI approaches are the norm at present and are primarily data-driven. Here, AI outputs are trained against known ground truths. Unsupervised AI, which is not reliant on label-based feedback, may be an attractive alternative. For this case, the network infers the internal structure of the input data, often through some policy of rewards for the outcome – as a result, the optical imaging takes on a form of 'intelligence' in the clinical decision-making process.

Role of AI in IVF

AI will likely demonstrate utility in several facets of the IVF procedure. Below, we discuss some of the more recent literature on the use of AI in gamete and embryo selection as well as the development of a treatment regimen (Fig. 2).

AI in gamete selection

Current clinical assessment of gamete health is focused on identifying early markers of quality. This includes visualisation of gametes either through direct inspection or via static images or time-lapse videos. The quality of female gametes is associated with follicle size, oocyte morphology, and cytoplasmic characteristics. In terms of sperm, morphology, concentration, and motility are known factors that are directly correlated with IVF success. However, the selection is prone to a high degree of variation between operators. AI can be transformative in this aspect. AI would remove the subjectivity of human assessment from the decision-making process, and objectively rank gametes based on quality. The use of AI for oocyte selection may be limited due to the practice of fertilising all available oocytes. Unless of course, AI was able to predict blastocyst formation or more importantly, live birth prior to fertilisation. Such capacity appears limited based on the current literature (Manna *et al.* 2013). The greatest benefit of AI may come from the selection of sperm for intracytoplasmic sperm injection (ICSI), a process currently performed by the embryologist. The development of new assessment criteria for sperm selection might arise through the use of unsupervised AI, where new markers of sperm quality are identified, such as swimming patterns, direction of motion, or difference in sperm compartments (i.e. length of head vs tail). Although more challenging, AI may compute the optimal sperm–egg combination to achieve the highest success rate or perhaps determine whether IVF or ICSI is the best fertilisation approach. Interestingly, AI has demonstrated proficiency in predicting fertility outcomes based on distinct ultrastructural details of mouse sperm – a larger sperm head compared to the midpiece is associated with improved blastocyst development (Kandel *et al.* 2020). Similarly, AI has yielded high accuracy in classifying human sperm using kinetic parameters – 89.9% (Goodson *et al.* 2017); as well as classifying sperm head morphology with a high concordance rate with current manual classifications (SCIAN and HuSHeM, 88 and 94%, respectively) (Iqbal *et al.* 2020). Notably, these studies used images or videos acquired during a routine assessment. AI can, therefore, complement current clinical practices whilst offering objective gamete selection to substantiate the assessment made by embryologists.

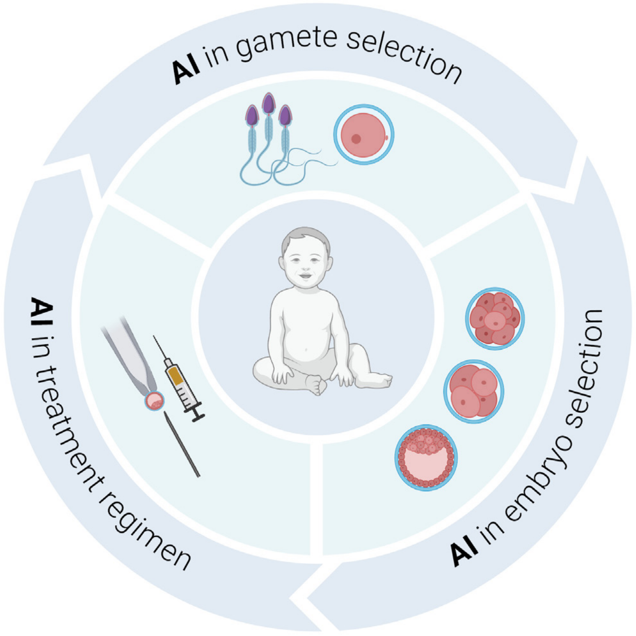

Figure 2 Emerging role of artificial intelligence (AI) in the *in vitro* fertilisation (IVF) clinic. AI represents an opportunity for technological advancement to improve IVF success. It is multifaceted in its capability. For example, AI may aid in selecting the best oocyte and sperm combination as well as predicting embryo quality. Furthermore, AI may assist the clinician in developing an optimal patient-specific treatment regimen to improve IVF success.

AI in embryo selection

Morphological assessment of embryos is the most commonly used process to select embryos for transfer. This occurs through direct visualisation using a light microscope or by time-lapse imaging. Both approaches grade embryos on their ability to reach particular stages of development in a timely manner. As with gamete selection, there is a high

https://doi.org/10.1530/RAF-21-0043

degree of variation between operators and clinics due to the subjective nature of these assessments. Consequently, standardisation is challenging within a clinic and near-impossible between institutions. As such, morphological grading remains limited in its ability to predict live birth outcomes (Chen *et al.* 2017). AI, using routinely generated images or time-lapse videos, may objectively and accurately grade and rank embryos, thus, assisting in the decision-making process to transfer or freeze them. Further, AI may have a role in analysing data from non-invasive metabolomic and secretory profiles from the embryo during culture. Consequently, this may lead to improved culture media formulations and regimens.

Studies using AI to predict embryo developmental competence are on the rise, as reviewed in Fernandez *et al.* (2020). These show promise in improving the accuracy of manual grading and predicting development to blastocyst and pregnancy (Chavez-Badiola *et al.* 2020, VerMilyea *et al.* 2020, Coticchio *et al.* 2021). Interestingly, AI is able to use cellular textures to accurately discern between the inner cell mass (fetal lineage, 91% accuracy) and trophectoderm (placental lineage, 86.6% accuracy) of the blastocyst (Saeedi *et al.* 2017). The use of AI to accurately identify cell lineages may lead to improved embryo classification and IVF success. However, whether AI can accurately predict the live birth outcome to a high degree, for example, in excess of 80–90% is yet to be demonstrated. However, a recent study has yielded a 69.3% accuracy (Vogiatzi *et al.* 2019). In summary, the implementation of AI in the embryo grading and selection process could provide an objective and accurate prediction of pregnancy and live birth outcomes.

AI in treatment regimen

In the IVF clinic, decision-making for an IVF cycle regimen is guided by patient age, gamete quality, medical history, and many more. This process intends to maximise the chances of pregnancy and birth of a healthy baby. From patient to patient, an IVF cycle might thus differ in stimulation protocol and mode of fertilisation (IVF vs intracytoplasmic sperm injection) as well as the potential for other procedures including assisted hatching and preimplantation genetic testing, amongst others. Planning for an IVF cycle is heavily reliant upon input from the clinician, who may prescribe a different treatment regimen based on their own clinical experience and/or clinic-to-clinic differences in training and in-house practice. AI could aid fertility practitioners in this aspect, enabling objective decision-making to optimise the treatment protocol for the best outcome. AI could also be applied to data-mine existing patient records to discover novel markers that predict pregnancy and live birth. Indeed, the use of AI in other fields such as oncology is very promising in this regard with a significant reduction in time required to formulate radiotherapy treatment plans (from days to minutes) (Wang *et al.* 2019). Interestingly, the use of AI to data-mine existing databases led to the identification of novel genes associated with the pathogenesis of endometriosis (Bouaziz *et al.* 2018). Incorporating AI within this aspect of the IVF cycle might lead to improved success as well as the reduced workload for clinicians.

Artificial intelligence: is there a limit?

Artificial intelligence promises a revolutionary change in the field of medicine, including the IVF clinic. There are, however, major challenges to overcome before AI can take centre-stage in IVF treatments. First is the onus of the ground truth or target for learning. To date, the vast majority of publications have utilised positive pregnancy test as the metric of success. It is our opinion that the most appropriate and patient-centred target is the birth of a healthy child. In time to come, this may shift to include the health of IVF-conceived children later in life. A further challenge facing learning AI is that they evolve with data, lacking transparency and interpretability to their inner workings – often approached with apprehension as 'black box' machines. Major advances in the past decade have enabled us to disentangle, guide, and evaluate AI. However, an explicit understanding of trained algorithms is unintuitive as, ultimately, we expect AI to supersede the capacity of ‘human intelligence’. AI mandates a regulatory and perhaps a mentality change in the clinic, wherein algorithms are not expected to be infallible but rather embracing the essential part of all ‘intelligence’ – the capacity to make and learn from mistakes. Aside from this, the required magnitude of data to train AI to achieve high accuracy and reliability is proportionally large to the extent of predictive accuracy and is also dependent on data quality. Therefore, AI may be ideal for clinics with an existing large patient database but may disadvantage others with smaller patient populations. This can be circumvented with a unified registry at a national or international level to minimise potential selection bias commonly seen with small data sets. On the flip side, the use of a unified AI and centralised data may predispose the electronic health system to potential data breaches or, potentially, may encourage some to optimise their digital information to influence AI outcomes. If AI generates a metric of quality

https://doi.org/10.1530/RAF-21-0043

Table 1 Steps toward adoption of AI in the IVF clinic.

Hurdles	Requirements to develop AI in the infertility service
• Most clinics are too small in isolation to have sufficient data for AI • Limited computing power and data storage • Discrepancy in recorded and reported patient characteristics and reproductive outcomes between clinics • Variation between studies in the input parameters used for AI to predict reproductive outcomes	• Consortia to develop standardised and detailed recording and reporting of patient data • Multi-clinic sharing of anonymised patient data • Increased computing power and data storage • Consensus on input parameters used in AI for predicting reproductive outcomes • Rigorous, high quality, multi-centre, peer-reviewed validation studies for AI algorithms

for a clinic, patients will likely seek out the best one. In an example of Goodhart's Law, clinics may try to optimise their 'AI score' by modifying the parameters that they report, or make them reluctant to treat 'riskier' patients in an effort to improve or retain their scores. Moreover, AI often requires re-training when switching from one sample to another, or when switching from one imaging modality to another, which can be both time-consuming and costly. Nonetheless, this may be overcome through unsupervised reinforcement learning, where the algorithm itself is able to learn how to extract features with optimal predictive accuracy for application in the medical field.

Future directions

Despite the challenges, it is undeniable that AI has a future role in the IVF field. However, AI must overcome hurdles prior to its utilisation (Table 1). AI requires not only extensive simulation studies and publication in peer-reviewed journals but also rigorous validation in numerous clinics and in different patient populations. More broadly in the clinical domain, amongst an extensive list of medical algorithms approved by the US Food & Drugs Administration (FDA) (Benjamens *et al.* 2020), only a handful of these are supported by peer-reviewed publications. It also highlights the need for more rigorous scrutiny and peer-reviewed evidence for future AI algorithm applications given the recent relaxation of FDA regulatory requirements for medical algorithm approval (Benjamens *et al.* 2020). Thus, we should evaluate AI not only by the capacity to enhance existing clinical assessments but on their ultimate improvement to real clinical targets, such as live birth. Nonetheless, it is an exciting journey that lies ahead. For the case of optimising an IVF cycle, AI will likely aid clinicians in devising patient-specific treatment regimens, before and after oocyte collection, as well as improving gamete and embryo selection. The integration of this new technology is well-positioned to improve IVF, promising an increase in the current and stagnant success rate. The continued emergence of AI studies will benefit the IVF industry and most importantly, patient's affected by infertility.

Conclusion

Incorporating AI technology into the IVF clinic may be the next frontier in the journey towards personalised reproductive medicine and improved fertility outcomes for patients. Infertility, as a global burden of disease, continues to rise (Sun *et al.* 2019). Furthermore, it is predicted that the global population will decline below replacement levels in the near future (Vollset *et al.* 2020). Additionally, the recent global COVID19 pandemic has delayed childbearing and fertility treatment for many. Thus, there will be an expected increase in demand for IVF moving forward. With the success of IVF stagnant for the past decade, AI is well placed to improve and enhance current clinical practices and the predictive power of IVF outcomes. This will undoubtedly benefit patients and ensure the next generation of healthy individuals conceived through IVF.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of this commentary.

Funding

K R D is supported by a Mid-Career Fellowship from the Hospital Research Foundation (C-MCF-58-2019). K D is supported by the UK Engineering and Physical Sciences Research Council (grants EP/P030017/1 and EP/R004854/1).

Author contribution statement

The commentary was planned jointly by D J X C, K D, and K R D. All authors wrote, reviewed and approved the final version.

https://doi.org/10.1530/RAF-21-0043

References

Benjamens S, Dhunnoo P & Mesko B 2020 The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digital Medicine* **3** 118. (https://doi.org/10.1038/s41746-020-00324-0)

Bouaziz J, Mashiach R, Cohen S, Kedem A, Baron A, Zajicek M, Feldman I, Seidman D & Soriano D 2018 How artificial intelligence can improve our understanding of the genes associated with endometriosis: natural language processing of the PubMed Database. *BioMed Research International* **2018** 6217812. (https://doi.org/10.1155/2018/6217812)

Chavez-Badiola A, Flores-Saiffe Farias A, Mendizabal-Ruiz G, Garcia-Sanchez R, Drakeley AJ & Garcia-Sandoval JP 2020 Predicting pregnancy test results after embryo transfer by image feature extraction and analysis using machine learning. *Scientific Reports* **10** 4394. (https://doi.org/10.1038/s41598-020-61357-9)

Chen M, Wei S, Hu J, Yuan J & Liu F 2017 Does time-lapse imaging have favorable results for embryo incubation and selection compared with conventional methods in clinical in vitro fertilization? A meta-analysis and systematic review of randomized controlled trials. *PLoS ONE* **12** e0178720. (https://doi.org/10.1371/journal.pone.0178720)

Coticchio G, Fiorentino G, Nicora G, Sciajno R, Cavalera F, Bellazzi R, Garagna S, Borini A & Zuccotti M 2021 Cytoplasmic movements of the early human embryo: imaging and artificial intelligence to predict blastocyst development. *Reproductive Biomedicine* **42** 521–528. (https://doi.org/10.1016/j.rbmo.2020.12.008)

De Geyter C, Wyns C, Calhaz-Jorge C, de Mouzon J, Ferraretti AP, Kupka M, Nyboe Andersen A, Nygren KG & Goossens V 2020 20 years of the European IVF-monitoring consortium registry: what have we learned? A comparison with registries from two other regions. *Human Reproduction* **35** 2832–2849. (https://doi.org/10.1093/humrep/deaa250)

Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM & Thrun S 2017 Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542** 115–118. (https://doi.org/10.1038/nature21056)

Fernandez EI, Ferreira AS, Cecilio MHM, Cheles DS, de Souza RCM, Nogueira MFG & Rocha JC 2020 Artificial intelligence in the IVF laboratory: overview through the application of different types of algorithms for the classification of reproductive data. *Journal of Assisted Reproduction and Genetics* **37** 2359–2376. (https://doi.org/10.1007/s10815-020-01881-9)

Goodson SG, White S, Stevans AM, Bhat S, Kao CY, Jaworski S, Marlowe TR, Kohlmeier M, McMillan L, Zeisel SH, *et al.* 2017 Casanova: a multiclass support vector machine model for the classification of human sperm motility patterns. *Biology of Reproduction* **97** 698–708. (https://doi.org/10.1093/biolre/iox120)

Iqbal I, Mustafa G & Ma J 2020 Deep learning-based morphological classification of human sperm heads. *Diagnostics* **10** 325. (https://doi.org/10.3390/diagnostics10050325)

Kandel ME, Rubessa M, He YR, Schreiber S, Meyers S, Matter Naves L, Sermersheim MK, Sell GS, Szewczyk MJ, Sobh N, *et al.* 2020 Reproductive outcomes predicted by phase imaging with computational specificity of spermatozoon ultrastructure. *PNAS* **117** 18302–18309. (https://doi.org/10.1073/pnas.2001754117)

Manna C, Nanni L, Lumini A & Pappalardo S 2013 Artificial intelligence techniques for embryo and oocyte classification. *Reproductive Biomedicine Online* **26** 42–49. (https://doi.org/10.1016/j.rbmo.2012.09.015)

Saeedi P, Yee D, Au J & Havelock J 2017 Automatic identification of human blastocyst components via texture. *IEEE Transactions on Bio-Medical Engineering* **64** 2968–2978. (https://doi.org/10.1109/TBME.2017.2759665)

Sun H, Gong TT, Jiang YT, Zhang S, Zhao YH & Wu QJ 2019 Global, regional, and national prevalence and disability-adjusted life-years for infertility in 195 countries and territories, 1990–2017: results from a Global Burden of Disease Study, 2017. *Aging* **11** 10952–10991. (https://doi.org/10.18632/aging.102497)

VerMilyea M, Hall JMM, Diakiw SM, Johnston A, Nguyen T, Perugini D, Miller A, Picou A, Murphy AP & Perugini M 2020 Development of an artificial intelligence-based assessment model for prediction of embryo viability using static images captured by optical light microscopy during IVF. *Human Reproduction* **35** 770–784. (https://doi.org/10.1093/humrep/deaa013)

Vogiatzi P, Pouliakis A & Siristatidis C 2019 An artificial neural network for the prediction of assisted reproduction outcome. *Journal of Assisted Reproduction and Genetics* **36** 1441–1448. (https://doi.org/10.1007/s10815-019-01498-7)

Vollset SE, Goren E, Yuan CW, Cao J, Smith AE, Hsiao T, Bisignano C, Azhar GS, Castro E, Chalek J, *et al.* 2020 Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study. *Lancet* **396** 1285–1306. (https://doi.org/10.1016/S0140-6736(20)30677-2)

Wang C, Zhu X, Hong JC & Zheng D 2019 Artificial intelligence in radiotherapy treatment planning: present and future. *Technology in Cancer Research and Treatment* **18** 1533033819873922. (https://doi.org/10.1177/1533033819873922)

Wijesinghe P & Dholakia K 2021 Emergent physics-informed design of deep learning for microscopy. *Journal of Physics: Photonics* **3** 021003. (https://doi.org/10.1088/2515-7647/abf02c)

Yamashita R, Long J, Longacre T, Peng L, Berry G, Martin B, Higgins J, Rubin DL & Shen J 2021 Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *Lancet: Oncology* **22** 132–141. (https://doi.org/10.1016/S1470-2045(20)30535-0)

Received in final form 16 August 2021
Accepted 23 August 2021