Artificial Intelligence in Reproductive Medicine – An Ethical Perspective

Künstliche Intelligenz in der Reproduktionsmedizin – eine ethische Betrachtung

OPEN ACCESS

Authors

Vasilija Rolfes[1], Uta Bittner[1,2], Helene Gerhards[2], Jan-Steffen Krüssel[3], Tanja Fehm[4], Robert Ranisch[5,6], Heiner Fangerau[1]

Affiliations

1 Institut für Geschichte, Theorie und Ethik der Medizin, Medizinische Fakultät, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany
2 Institut für Sozialforschung und Technikfolgenabschätzung, Ostbayerische Technische Hochschule Regensburg, Regensburg, Germany
3 Klinik für Frauenheilkunde und Geburtshilfe, Universitäres interdisziplinäres Kinderwunschzentrum Düsseldorf, Medizinische Fakultät, Heinrich-Heine-Universität Düsseldorf, Universitätsklinikum Düsseldorf, Düsseldorf, Germany
4 Klinik für Frauenheilkunde und Geburtshilfe, Universitätsklinikum Düsseldorf, Medizinische Fakultät, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany
5 Juniorprofessur für Medizinische Ethik mit Schwerpunkt auf Digitalisierung, Universität Potsdam, Fakultät für Gesundheitswissenschaften Brandenburg, Potsdam, Germany
6 Forschungsstelle „Ethik der Genom-Editierung“, Institut für Ethik und Geschichte der Medizin, Eberhard-Karls-Universität Tübingen Medizinische Fakultät, Tübingen, Germany

Key words

artificial intelligence, reproductive medicine, infertility, medical ethics, research ethics

Schlüsselwörter

künstliche Intelligenz, Reproduktionsmedizin, Infertilität, Medizinethik, Forschungsethik

received 23.12.2021
accepted after revision 29.5.2022

Bibliography

Geburtsh Frauenheilk 2023; 83: 106–115
DOI 10.1055/a-1866-2792
ISSN 0016-5751

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Correspondence

Vasilija Rolfes
Institut für Geschichte, Theorie und Ethik der Medizin, Medizinische Fakultät
Heinrich-Heine-Universität Düsseldorf
Moorenstraße 5
40225 Düsseldorf, Germany
Vasilija.Rolfes@uni-duesseldorf.de

Deutsche Version unter:
https://doi.org/10.1055/a-1866-2792.

ABSTRACT

Artificial intelligence is steadily being integrated into all areas of medicine. In reproductive medicine, artificial intelligence methods can be utilized to improve the selection and prediction of sperm cells, oocytes, and embryos and to generate better predictive models for in vitro fertilization. The use of artificial intelligence in this field is justified by the suffering of persons or couples who wish to have children but are unable to conceive. However, research into the use of artificial intelligence in reproductive medicine is still in the early experimental stage and furthermore raises complex normative questions. There are ethical research challenges because evidence of the efficacy of certain pertinent systems is often lacking and because of the increased difficulty of ensuring informed consent on the part of the affected persons. Other ethically relevant issues include the potential risks for offspring and the difficulty of providing sufficient information. The opportunity to fulfill the desire to have children affects the welfare of patients and their reproductive autonomy. Ultimately, ensuring more accurate predictions and allowing physicians to devote more time to their patients will have a positive effect. Nevertheless, clinicians must be able to process patient data conscientiously. When using artificial intelligence, numerous actors are involved in making the diagnosis and deciding on the appropriate therapy, raising questions about who is ultimately responsible when mistakes occur. Questions of fairness arise with regard to resource allocation and cost reimbursement. Thus, before implementing artificial intelligence in clinical practice, it is necessary to critically examine the quantity and quality of the data used and to address issues of transparency. In the medium and long term, it would be necessary to confront the undesirable impact and social dynamics that may accompany the use of artificial intelligence in reproductive medicine.

ZUSAMMENFASSUNG

Künstliche Intelligenz findet Eingang in alle Bereiche der Medizin. In der Reproduktionsmedizin können Methoden der künstlichen Intelligenz dazu genutzt werden, die Auswahl und Prädiktion von Spermazellen, Eizellen und Embryonen zu verbessern sowie bessere Vorhersagemodelle für die In-vitro-Fertilisation zu generieren. Der Einsatz künstlicher Intelligenz wird durch das Leiden der Personen oder Paare an einem unerfüllten Kinderwunsch gerechtfertigt. Die Forschung zum Einsatz von künstlicher Intelligenz in der Reproduktionsmedizin befindet sich aber noch in einem sehr frühen experimentellen Stadium und wirft zudem komplexe normative Fragen auf. Forschungsethische Herausforderungen bestehen aufgrund der häufig fehlenden Evidenzen bezüglich der Wirksamkeit entsprechender Systeme sowie aufgrund der erschwerten Bedingungen einer informierten Einwilligung der Betroffenen. Ethisch relevant sind zudem mögliche Risiken für die Nachkommen und die adäquate Aufklärung. Die Chance, einen Kinderwunsch zu erfüllen, wiederum berührt das Patient*innenwohl und den Ermöglichungscharakter der reproduktiven Autonomie. Zuletzt haben bessere Prognosen und mehr Zeit der Ärzt*innen für ihre Patient*innen einen positiven Effekt. Gleichwohl müssen Kliniker*innen mit den Patient*innendaten verantwortungsbewusst umgehen können. Beim Einsatz von künstlicher Intelligenz ist eine Vielzahl an Akteur*innen bei der Diagnose- und Therapiestellung beteiligt, woraus sich Fragen nach Verantwortlichkeiten bei Fehlern ergeben. Gerechtigkeitsfragen stellen sich bezüglich der zukünftigen Kostenerstattung und Ressourcenallokation. Bis zum klinischen Einsatz der künstlichen Intelligenz sind zuletzt die Datenquantität und -qualität kritisch zu betrachten und Fragen der Transparenz zu lösen. Mittel- und langfristig wäre eine Auseinandersetzung mit unerwünschten Effekten und Sozialdynamiken notwendig, die mit dem Einsatz von künstlicher Intelligenz in der Reproduktionsmedizin einhergehen könnten.

Introduction

Artificial intelligence (AI) is developing rapidly, and AI applications are being implemented in various sectors, including several medical fields [1, 2]. In medicine, complex AI algorithms can be used to analyze large amounts of data with the aim of improving diagnoses, prognoses and preventive measures [3]. In recent years, image data analysis has become a promising area of application, and it has even been suggested that the results generated by AI are superior to the contributions of experts [4]. Moreover, AI is expected to increase the efficacy of workflow processes in hospitals and improve patient monitoring [5].

In reproductive medicine, new areas of application of AI-based methods are being explored. Persons or couples who desire to have children but cannot conceive are often in a life crisis and have a reduced quality of life [6]. An inability to conceive may be viewed as an obstacle that frustrates the generally accepted need to reproduce and care for a child; this may acquire a psycho-existential dimension, as the inability to have children threatens people's existential vision of their future [7]. Over 40 years ago, the birth of Louise Brown marked the start of a new era that offered a new hope to infertile persons and couples, as the possibility of achieving a successful pregnancy via in vitro fertilization (IVF) became a reality [8]. According to some estimates, up to 186 million people worldwide are affected by infertility [9]. The continued development and improvement of reproductive technologies, which include the cryopreservation of oocytes and embryos, IVF, preimplantation genetic diagnostics (PGD), and the technological option to select preimplantation embryos based on morphokinetic criteria have greatly increased clinical pregnancy rates over the past 40 years [10].

However, several challenges persist. On average, the current probability in Germany of giving birth after a fresh embryo transfer following IVF or intracytoplasmic sperm injection (ICSI) is approximately 24%; the probability when using previously cryopreserved and thawed pronuclear-stage or multicellular-stage embryos is 20%. In essence, despite successful fertilization and cell division only every fourth fresh embryo transfer and every fifth cryo-transfer will result in the live birth of an infant [11]. The age of the oocytes of the woman desiring to have children and the quality of the embryos are the most crucial factors determining whether an IVF treatment would result in pregnancy [11, 12]. Reliable methods that allow a more precise assessment of the quality of oocytes, sperm, and embryos are lacking. While the option of using preimplantation diagnostics prior transferring the embryo to the uterus exists, this approach is ethically controversial. In Germany, it is legal only within certain limits, furthermore, it is technically complex, expensive, and poses the risk of injuring or destroying the embryo [13, 14].

The driving force behind the development of AI-based technologies in reproductive medicine is the intent to improve the treatment and prognosis of infertile patients by taking large amounts of data and combining them to obtain meaningful results [14, 15]. It is in this context that our ethical discussion is relevant. The discussion is framed in part by the needs of persons with an unfulfilled desire to have children, and in part by the use of specific AI technologies to achieve medical goals. The principles of healing, helping, and alleviating, which can be subsumed under the general principle of beneficience, provide the ethical framework for this discussion [16]. Extracorporeal fertilization and selective reproduction (e.g., after PGD) are still considered controversial [17]. Research and the potential integration of AI into clinical care raise complex normative questions with regard to content and procedural issues. Thus, a proactive ethical debate is necessary, even if AI research and the use of AI in reproductive medicine are still in very early experimental stages.

Given this context, we will start by providing a brief outline of the current status of the research and development in this field, followed by a structured ethical analysis of the potential uses of AI-based methods in reproductive medicine. After that, we will discuss the impact of the relevant technologies on the relationship between physicians and their patients. In the conclusion, we will consider possible future developments in AI-driven reproductive medicine and briefly address possible social challenges associated with these trends.

Possible Applications of AI in Reproductive Medicine

In Europe, pregnancy rates and live birth rates for all treatment options vary widely between countries. In 2015, the pregnancy rate following fresh cycles after IVF or ICSI was between 19.6 % and 44 %, and live birth rates ranged from 10.2 % to 40 %. The rate of live births after the transfer of frozen-thawed embryos ranged between 12.8 % and 37.5 % in different European countries [18]. Overall success rates decline sharply as women become older [19]. A retrospective analysis by Stoop et al. of live births after fresh and cryopreserved embryo transfers showed that the average rate per harvested ripe oocyte was 4.47% for women between the ages of 23 and 37 [20]. Therefore, it can be assumed that assisted reproductive medicine could potentially be improved in several areas. Now that AI systems are coming into more general use, it is hoped that the automatic classification of sperm, embryos, and oocytes will be possible, thereby increasing the success rates of IVF [14].

Basics of AI

The definition of AI in the context of this text is as following: using complex algorithms to imitate logical thinking and cognitive functions. Several heterogeneous techniques are included under the umbrella term AI. Machine learning (ML) is a particularly successful AI application. ML identifies patterns between variables in large datasets. Previously unknown correlations can be identified with ML and used to generate new hypotheses and pioneer new areas of research [21]. Most ML approaches can be classified as either supervised or unsupervised approaches. With supervised ML, labelled training data are used to develop models in which the target result (e.g., a diagnosis) is known. In contrast, unsupervised ML, does not require labeled data. Instead, patterns or aggregations which occur within the data are recognized [22]. Deep learning (DL) is a variant of ML. DL attempts to imitate the functions of the human brain, using different levels of artificial neural networks to generate automatic predictions based on training datasets [21].

Despite several potential pitfalls, a promising clinical approach involves making decisions for infertility patients based on an analysis of a variety of medical data. Reproductive medicine specialists can use ML models to help identify suitable treatments for persons struggling with infertility [22]. We present concrete approaches below.

Selection and prediction of sperm cells

Sperm morphology is a common cause of (partial) infertility in men. Computer-assisted sperm analysis systems are already in use. Nevertheless, the analysis of sperm motility remains difficult due to the clumping of spermatozoa and other contributing factors [23]. Moreover, analyses may vary between laboratories. Finally, in approximately one third of men, no clear aetiology can be identified [24], which means that the cause of their infertility cannot be identified with standard methods [14]. In the future, automatic methods based on image may help obtaining more objective and precise results [14]. Goodson et al., for example, used retrospective data from 425 human spermiograms to develop models which identified chromosomal anomalies. Size, total testicular volume, follicle-stimulating hormone, luteinizing hormone, total testosterone and ejaculate volume were used as the input data; based on these data, the predictive accuracy for chromosomal anomalies was over 95% [25]. Using data-mining methods, a research team also developed two specific artificial neural networks to predict human sperm concentrations and motility according to environmental factors and the men's lifestyles [26].

Evaluation and selection of oocytes

Successful reproduction, whether spontaneous or assisted, depends significantly on the quality of the oocytes. However, the mechanisms of embryo malformations that develop from oocytes of insufficient quality are not known [27]. Yanez et al. attempted to predict the developmental potential of human oocytes by measuring the viscoelastic properties of human zygotes just a few hours after fertilization without destroying the zygote. They were able to reliably predict viability and blastocyst formation with an accuracy of over 90%, a specificity of 95% and a sensitivity of 75%. The researchers also investigated RNA-sequencing data using a *support-vector-machine classifier* and found that non-viable embryos had significantly different transcriptomes, particularly with regard to the expression of genes important for oocyte maturation [27]. Cavalera et al. observed mouse oocytes during in vitro maturation and took pictures for time-lapse analysis; the obtained data were analyzed using artificial neural networks. They were able to identify oocytes capable or incapable of further development with an accuracy of 91.03% [28].

Evaluation of embryo quality

Saeedi et al. presented the first automatic method to segment two main components of human blastocysts, the trophectoderm (i.e., the external cell layer of the blastocyst between the 4^{th} and 7^{th} day after fertilization), and the inner cell mass. The two regions are strongly textured and the quality of their textures is quite similar. On imaging, they often look as though they are connected to each other. The automatic identification of both regions facilitates the detailed evaluation of blastocysts. Saeedi et al. reported an accuracy of 86.6% for the identification of the trophectoderm and 91.3% for the inner cell mass. The aim of their study was to achieve a better understanding of why the transfer of certain embryos results in higher pregnancy rates than the transfer of other embryos [29]. In a study published in 2019, researchers from Cornell University trained a deep-learning algorithm from Google to recognize embryos with high, moderate, or low development potential, measured according to the probability of them successfully implanting after intrauterine transfer. To train the artificial neural networks, they used more than 10000 time-lapse images of human embryos. They additionally developed a decision tree using the clinical data of 2182 embryos, which combined the quality of the embryo with the age of the patient and defined scenarios which are associated with the probability of becoming pregnant. Their analysis showed that the probability of becoming pregnant following the transfer of an individual embryo ranged between 13.8% (age ≥ 41 and poor quality) and 66.3% (age < 37 and good quality); this correlated with the results of the automatic evaluation of blastocyst quality and the age of the patient [30].

Predictive models for IVF

Already in 1997, Kaufmann et al. developed a model to predict the probability of successful IVF. For this, they used an artificial neural network with a predictive value of 59%. The research group used four criteria: age of the intended mother, number of oocytes obtained, number of transferred embryos, and whether the embryos had previously been cryopreserved [31]. In a similar study conducted in 2010, the datasets of 250 patients were collected from IVF research centers, fertility clinics and maternity hospitals. To train and test the algorithm, the research group used significantly more criteria (age of the woman, duration of infertility, body mass index [BMI], previous pregnancies, previous operations, endometriosis, fallopian tube problems, ovulation factor, sperm concentration, sperm vitality, number of aspirated oocytes, number of transferred embryos, prior history of miscarriage, and psychological factors) and achieved an accuracy of 73% for their results [32].

In a recent retrospective study, the researchers aimed to determine whether a simple prognostic algorithm could differentiate between couples who require treatment for infertility and couples who can initially be offered less invasive strategies, the couples were divided into groups according to their medical need for IVF treatment and their prognosis for achieving a natural conception. A Kaplan-Meier curve was generated for each group to measure the probability of conceiving naturally over time and the effect of infertility treatment. The outcome was that for couples with slight or unexplained male infertility, the chance of live birth without treatment and poor prognosis increases significantly, in comparison to couples with a good prognosis. This prediction model provides the opportunity to individualize fertility treatments, thus avoiding unnecessary IVF treatments without affecting the chances of fertility. [33].

Ethical Aspects

A discussion of the ethical aspects implicit in the use of AI in reproductive medicine is presented below. We found that the ethical aspects can be divided into four main areas, as shown in the table (▶ **Table 1**):

Research ethics

As shown by examples from reproductive medicine research, using AI in medicine opens up new possibilities in clinical applications. However, up to now, even promising developments have rarely been brought to technical maturity, and even fewer have become part of standard clinical practice. In ideal conditions, the methods of medical ML in particular could be highly effective, for example for analysing imaging data. But performance in clinical practice is considerably poorer, which might also be the case for AI-supported imaging analysis of spermiograms. This is unsurprising, as the overwhelming majority of studies with AI or ML in medical research are performed in a retrospective setting [34]. Investigations into AI applications also have indicated that the study designs have limitations, creating further difficulty making statements about their effectiveness in the clinical setting [35]. Recently, the dearth of randomized controlled studies (RCT) has led to some criticisms about the validity of existing research and development in the field of medical AI and ML [36].

In recent decades, developments in reproductive medicine such as PGD or ICSI were often transferred directly from laboratories to clinical applications without undergoing a comprehensive review of their effectiveness and safety [37]. Studies in the field of reproductive medicine conducted in humans are particularly sensitive from the standpoint of research ethics and also present very particular challenges. While initially, the intended mother is the person who is directly affected by the study (and, in contrast to her male partner, also bears the majority of physical and psychological burdens), research and expected innovations in the field of assisted reproduction (e.g., AI applications) can result in the birth of offspring. The offspring are inevitably affected by the possible risks of experiments, despite not having had the option to consent to participating in the study. Similarly, potential long-term effects on offspring caused by an experimental procedure performed during assisted reproduction may possibly only be fully understood several years later. This problem is compounded by the ethical questions that arise during the testing of new procedures based on the use or non-use of human embryos.

▶**Table 1** Ethical aspects of AI in reproductive medicine.

	Potential opportunities	Potential risks
Research ethics	1. Developing improved treatments for infertility treatment options → increased probability of reproduction	1. Difficulties to adequately inform patients/study participants' 2. Long-term monitoring not always possible 3. Unrealistic hopes of study participants 4. Moral status of human embryos
Impact on the chances and autonomy of patients	1. Higher baby-take-home rate 2. Lower therapy discontinuation rates as well as lower physical and psychological stresses 3. Social stigma of childlessness avoided 4. Support of reproductive autonomy	1. Lack of evidence on effectiveness 2. Insufficient information given to patients/study participants
Physician-patient relationship	1. Personalized information for patients 2. Better therapy options to treat infertility 3. Increases time resources in clinical practice	1. Informing patients is challenging because of the complexity of algorithms 2. Treatment results for the individual patient are unclear 3. Responsibilities, transparency and trust still not clear
Reproductive justice and chances of access	1. Lower financial burden 2. Creates a hierarchy of patients according to the probability of success	1. No widespread implementation 2. Follow-up costs due to complex technical operation (e.g., maintenance; liability) 3. Access to specialized "repro AI centres"

In all future research into the use of AI in reproductive medicine, providing well-considered information to the intended mother or intended parents will be crucial. It is also important to be aware of unfavourable situations that may arise when providing information to potential mothers or parents: study participants may nurse false hopes or unrealistic expectations when modern, attention-grabbing innovations are being tested. Especially women or couples with a longstanding painful experience of childlessness must be considered as a vulnerable group. As those participating in such studies may feel like this is their "last chance" to fulfil their desire to have a child, special precautions are necessary to ensure that participation is voluntary.

Patients' well-being and autonomy

One of the stated goals of using AI systems in predictive models for IVF is to improve outcomes compared to results obtained using conventional reproductive medicine methods. One measure of such improvements would be a higher baby-take-home rate, i.e., a higher probability that assisted fertilisation will result in a live birth. This could reduce the psychological and the physical suffering of patients. The normative principle of the primacy of the patient's welfare in assisted reproductive medicine includes providing the best-possible suitable treatment, based on objectively measurable medical parameters, and the patient's subjective experience, which involves taking into account of the patient's preferences in the treatment setting and the patient's satisfaction with the treatment. The use of AI technology could, in future, offer benefits in both areas. An analysis of 122 560 treatment cycles in Germany showed that 45 699 patients discontinued therapy after the birth of a child. The remaining 76 861 (62.7 %) patients discontinued therapy before they were able to fulfil their desire to have children [38]. A variety of reasons were given for discontinuing therapy: the absence of transferable embryos due to immature oocytes, the inability to harvest oocytes, failed fertilization attempts, or arrested embryo development. Such factors can discourage patients. Other reasons included no or insufficient response to stimulation, overstimulation syndrome or premature ovulation as well as incorrect administration of hormone injections leading to an unsuccessful course of treatment. Failure to conceive despite undergoing many reproductive medical procedures over a long time is also believed to be one of the reasons why patients discontinue therapy [38]. Other studies have come to similar conclusions and, in addition to these physical and psychological burdens, also have mentioned relationship difficulties and other personal problems as potential reasons for discontinuing treatment [39]. Furthermore, infertility is often tainted by social stigmas such as shame and social exclusion [40, 41]. More effective and faster treatment, which is conceivable in future AI applications, could provide some couples with a technological option to avoid these stresses, at least in partially, and thereby contributing to their well-being.

The use of AI in reproductive medical practice can also be analyzed in the context of patients' reproductive autonomy [42, 43]. Reproductive autonomy is a normative concept and should be understood as the capacity of individuals to decide freely, well-informed and without interference from others about their own reproduction. From this perspective, measures that support and enable patients to exercise their reproductive freedom such as the use of AI in reproductive medicine should ideally be available to everyone who want to have children but cannot [44]. Conversely, limiting reproductive autonomy should only be permissible if the use of new technologies in reproductive medicine would demonstrably result in an harm to the patient or their potential offspring [45]. Reproductive autonomy can also be interpreted as a right of entitlement, because being able to actuate one's own reproduction or non-reproduction can be viewed as a central aspect of a person's identity. Thus, it could be concluded that attempts must be made to support couples fulfilling their desire to have children [44]. This includes, for example, ensuring that access to and use of future reproductive technologies including AI applications is granted to such couples. However, the risks posed by an unfavourable information situation as outlined in the previous chapter must be avoided to allow patients to fully express their reproductive autonomy in this still experimental field.

Benefits and challenges for the physician–patient relationship

The potential clinical implementation of AI in reproductive medicine will inevitably impact on a lynchpin of the medical practice: the relationship between physician and patient. The use of AI may require reproductive medicine specialists to rethink their professional role, as they are the bridge between the algorithmic output and treatment-relevant decisions [46]. The physician must not only address the biological factors behind the patient's infertility but also patients' particular psychosocial and emotional stresses during treatment, which are well-known and have been demonstrated in studies [47]. In this context, the future use of AI could be beneficial as it will be possible to predict the success rate of individual patients better in terms of their likelihood of becoming pregnant. There is existing data that can be used to make probability statements, for example, based on the age of the woman [48]. However, an optimized and more precise prognosis could provide physicians with the opportunity to ensure patients are better informed and provide appropriate therapy recommendations. An option to optimize the selection of sperm, oocytes, and embryos would allow the treating physician to offer better and more efficient treatment. While the contribution of potential AI applications to such improvements may be minimal, given the stresses of infertility treatment, even those could offer valuable benefits. Moreover, it is hoped that the use of automated support systems in medicine will allow physicians to devote more time to the physician–patient relationship [49].

The potential implementation of AI into reproductive medicine would also place demands on medical staff. On the one hand, they are responsible for collecting and recording personal data including age, weight, and lifestyle information, etc., which will make it possible to train algorithms that could help facilitate pregnancies in the future. However, reproductive medicine specialists must also explain the use of algorithmic decision support systems to their patients and provide them with appropriate information: How can using AI technology have a more positive impact on the respective diagnosis and/or treatment? Is the physician convinced of the benefit of the AI support or not, and why? [50] On the other hand, the medical staff treating patients must ensure that the prognosis, diagnosis or treatment recommendation supported by AI systems do not contravene medical state-of-the-art and their professional judgment [50].

If future AI systems are going to place a greater emphasis on the quality of the gametes and determine their quality, the focus could move away from the individual patient. Predictive analysis models and the resulting treatment recommendations which are based on large amounts of data may be able to improve the treatment results for specific patient cohorts but they may not be necessarily beneficial for an individual patient. Consequently, these circumstances may come into conflict with the physician's obligation to act in the best interests of every individual patient [51]. Over the course of the various examinations, data collection procedures and analysis, there is a considerable risk that patients could become mere "data subjects" [3] and may not be perceived as persons [3]. Treating physicians must be aware of the potential dynamics involved in the datafication of people and continue to pay attention to the individual patient.

Questions of liability are another problematic area associated with the use of AI. With the increasing digitization of medicine and the use of ML algorithms, new players are increasingly entering the healthcare system [52]. These include tech companies and programmers, who play an important part in developing, training and testing ML systems. If AI applications result in treatment errors and wrong diagnoses, this will raise new questions regarding who is responsible [53]. This problem is compounded by how ML applications can appear to be a type of black box [54]. At time algorithms with a high validity can no longer be explained or the explanations would entail considerable effort or expense. The opacity of AI applications can make medical decision-making more difficult for specialists, as it may be unclear when they can rely on automated systems. The lack of transparency may also lessen patients' trust in relevant AI applications [55]. In addition, human–machine interactions pose certain challenges. Physicians with extensive experience in their field appear to have a greater mistrust of AI systems, while less experienced physicians may place excessive confidence in such systems [56].

Reproductive justice, and access

In Europe, the regulations on the reimbursement of the costs of assisted reproduction are very heterogeneous. Criteria regulating access to reproductive technology, such as age (both of the woman and the man intending to have children), whether the intending parent already has children, or how many treatment cycles the couple or the woman has already had, can differ between countries. In some countries, even female BMI can be a criterion for receiving public funds [57]. There are also considerable differences between countries with regard to the three main cost areas: medication, personnel, and laboratories. In some countries it is also relevant whether the IVF center is publicly funded or a private facility [57]. Hence infertile persons' options for accessing assisted reproductive technology may differ, resulting in social inequalities [58]. In Germany, for example, only 50% of the costs of a maximum of three treatment cycles with IVF or ICSI are covered for persons insured under the German public health insurance scheme, which consists of approximately 90% of the total population. Moreover, there are age limits that further restrict claims for reimbursement: the woman must be between the age of 25 and 39 years and the man between 25 and 49. Given this background, it would appear that access to assisted reproduction differs significantly from access to other healthcare services and that it is closely linked to cultural or moral norms and ideas of justice [59].

The prospect that AI procedures could soon be ready for use in clinical practice raises the hope that it will be possible to offer infertile patients more efficient therapy, which could lead to successful pregnancies and the reduction of the financial burden. However, in addition to these possible positive aspects, it is also important to be aware of other implications. Reducing the costs for individual patients and for the collective payor will only occur if the use of the new technology is efficient and if the AI-based technologies do not involve (disproportionately) higher costs generated by various factors, such as purchase, operation, data processing and storage, maintenance and updating of the model, visualization, the need for skilled operators, rectification of mistakes, possible liability costs, etc. in clinical practice [60]. Furthermore, considering the problem of fairness raises questions about access or barriers to access: while it is debatable whether AI-based systems will soon become part of clinical practice, it can be assumed that they will not be widely available immediately. It is conceivable that only a few reproductive medicine centres will initially include these in their list of services. Patients who do not have access to these centres may have to confront the likelihood their chances of success being lower, and reproductive medicine facilities that are unable to offer these services and do not provide the possible new "gold standard" could suffer a comparative disadvantage in terms of demand for their services. A concentration of AI-based reproductive medicine services in high-tech centres may cause the costs of procedures and measures that are considered necessary or desirable to initially increase. While this dynamic response to the use of AI may represent an advantage for the reproductive autonomy of some individuals, from the perspectives of reproductive justice and of reproductive medicine as a market economy, where healthcare services and costs are financed collectively, this concentration can represent a disadvantage for certain groups of patients with insufficient financial resources or mobility [61]. Hence, when conducted ethical assessments and weighing the consequences of technological change, attention must also be paid to potential barriers to access, issues of availability, and the financing of services.

A further aspect of justice concerns not only the question of costs but also the practical implementation in hospital. Based on critical reflections on the increasing "quantification of the social fabric" and its associated effects [62] it should be considered whether the seemly more precise measurement of the success rates of pregnancy could lead to a hierarchical categorization of persons requiring treatment and thereby engendering additional inequalities. The literature on the (potential) use of algorithms has raised concerns that algorithms could reproduce existing inequalities through the persons who design the algorithms or the data used to train the algorithms [63, 64]. When weighing the (potential) use of AI in infertility treatments, it is important to consider the question of which data are being used to train the algorithms, ideally to prevent discrimination against certain groups of patients.

Discussion

As in various other medical fields, there are several obstacles and risks associated with the potential use of supporting AI systems in reproductive medicine. It should be noted that AI-based methods in reproductive medicine are still in the early stages of development, making it very difficult to weigh what the actual risks and opportunities could entail. Furthermore, ML models can often not be fully explained and may be perceived as a "black box" [65]. This could result in a certain scepticism among clinicians and patients with regard to the diagnoses and therapy recommendations. The new scenario could result in patients feeling helpless when confronted with the use of non-transparent tools and automated decision-making processes that affect important aspects of their personal lives [50], creating additional uncertainty in a group of patients who cannot have children and are already physically and psychologically vulnerable. In contrast, for physicians, the opacity of AI applications may result in over-reliance on such systems or excessive distrust of them. Both scenarios could disadvantage patients. On the one hand unrealistic expectations about the efficiency of new technologies may give patients unjustified hope regarding their desire to have children. On the other hand, unwarranted scepticism about a highly effective AI system may result in the potential of such innovations being neglected.

According to recent studies, there are several limitations with regard to the quantity and quality of data, which significantly influence the performance, applicability and generalizability of the trained model. The majority of studies in the field of reproductive medicine have small sample sizes and are retrospective. Largescale randomised controlled studies which could test the validity of the algorithms and optimize their utilization are still lacking. Hence more research into personalized diagnosis and treatment, medical expert systems, and AI-supported reproduction is necessary [14].

Supervised ML and unsupervised ML raise the question of what physicians can know or say about the use of AI to affect the outcomes of infertility treatments. Physicians are obliged to handle with AI-generated recommendations even though they occasionally may not entirely understand the system and/or agree with the system's recommendations regarding the diagnosis and therapy [50]. This could make it more difficult to provide transparent and patient-focused information when using AI-supported data processing [66].

Overall, it appears that it will be necessary to examine the moral hazards described above in more detail from the standpoint of empirical ethics [67]. In such a context, the focus would not be on making an "objective" assessment and weighing the risks of different AI applications in reproductive medicine; rather would be on the (subjectively conveyed) conditions that affect the desirability, acceptance or rejection of AI in reproductive medicine. If these challenging areas are addressed in the course of qualitative interviews with physicians, persons with a desire to have children, AI researchers, providers and other groups, additional insights into the expected benefits and perceived risks of AI in reproductive medicine could be gained from the perspective of the affected persons before potential applications are integrated into standard care.

Conclusion: Future Direction of AI Applications in Reproductive Medicine and an Ethical Assessment

Improving the capability and capacity of AI technologies over time and integrating them into the treatment process could benefit patients and physicians by making high-quality reproductive medicine more effective and precise and by providing support to the treating physicians during decision-making. The future utilization of AI in reproductive medicine prior the starting IVF treatment could offer promising results and allow better prognoses to be made with regard to treatment success. This would give persons with an unfulfilled desire to have children the opportunity to address their individual chances of having a child early on with the aid of reproductive medicine specialists.

However, developments in the field of reproductive medicine must also be examined to assess the extent to which they could promote undesirable medium- and long-term effects and social dynamics. At the global level, reproductive medicine is not merely a growing research area but also a lucrative industry in which many different actors are vying for the attention of potential parents and clinics. In the context of popular AI applications in particular, innovations promise new feasibilities, whether in prediction of treatment outcomes or prenatal diagnostics. There are already companies that offer algorithm-supported embryo selection even for polygenic traits, for example, to exclude the risk of schizophrenia or cancer in offspring [68]. While the scientific basis of such services remains insufficient, this will create demand and generate further optimization fantasies.

It also remains to be seen to what extent the existing processes and logics underpinning the quantification of life phenomena could emerge or be reinforced by the use of AI-supported decision-making systems in reproductive medical practice, for example, using AI systems to reduce complex decisions to their metric values, resulting in a binary differentiation between "good" and "bad" embryos. According to Mau, descriptions represented in numerical units always express "values attributions", comparisons and "value orders" [62] and hence are hardly unbiased. They mirror attitudes and social effects such as acceptance and non-acceptance, and may thus steering behaviours [69]. For instance, it has become apparent over several years that countries that offer systematic prenatal screening for trisomy 21 have higher abortion rates of such pregnancies [70].

If decisions in fertility clinics will, in the future, increasingly be made with machine support, it will be necessary to critically observe and examine whether and to what extent pre-programmed value categories are the defining standards that mold the opinions of physicians as well as potential parents (cf. [63]). A machine calibrated to optimize outcomes may ultimately even become the pacemaker of a new form of eugenics, even if no identifiable "eugenicists" are involved [71]. This makes it all the more essential to ensure that the values underpinning the algorithms are transparent and that they are discussed publicly. This would help AI-supported medicine become a humane medicine.

Fundings

| Jürgen Manchot Stiftung (VR, UB) |

Acknowledgement

This article was partly conceived in the context of the project "Decision-making with the help of artificial intelligence methods" (Vasilija Rolfes, Uta Bittner) supported by the Jürgen Manchot Foundation. The authors would like to thank the anonymous reviewers for their helpful comments on the earlier version of this article. / Der vorliegende Beitrag ist zum Teil in dem von der Jürgen Manchot Stiftung geförderten Projekt „Entscheidungsfindung mit Hilfe von Methoden der Künstlichen Intelligenz" (Vasilija Rolfes, Uta Bittner) entstanden. Die Autorinnen und Autoren danken den anonymen Reviewern für die hilfreichen Kommentare zu der früheren Version dieses Artikels.

Conflict of Interest

The authors declare that they have no conflict of interest.

References/Literatur

[1] Gomez-Gonzalez E, Gomez E, Rivas-Marquez J et al. Artificial intelligence in medicine and healthcare: a review and classification of current and near-future applications and their ethical and social impact. arXiv 2020. doi:10.48550/arXiv.2001.09778

[2] Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019; 6: 94–98. doi:10.7861/futurehosp.6-2-94

[3] Jannes M, Friele M, Jannes C, Woopen C. Algorithmen in der digitalen Gesundheitsversorgung. Eine interdisziplinäre Analyse. Gütersloh: Bertelsmann Stiftung; 2018. . doi:10.11586/2019053

[4] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019; 25: 44–56. doi:10.1038/s41591-018-0300-7

[5] Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. Gastrointest Endosc 2020; 92: 807–812. doi:10.1016/j.gie.2020.06.040

[6] Sexty R, Griesinger G, Kayser J et al. Fertilitätsbezogene Lebensqualität bei Patientinnen in deutschen Kinderwunschzentren. Geburtshilfe Frauenheilkd 2016; 76: V14. doi:10.1055/s-0036-1571376

[7] Westermann AM, Alkatout I. Ist unerfüllter Kinderwunsch ein Leiden? – Der Leidensbegriff im Kontext der Kinderwunschtherapie. Ethik Med 2020; 32: 125–139. doi:10.1007/s00481-019-00556-z

[8] Steptoe PC, Edwards RG. Birth after the reimplantation of a human embryo. Lancet 1978; 312: 366. doi:10.1016/s0140-6736(78)92957-4

[9] Inhorn MC, Patrizio P. Infertility around the globe: new thinking on gender, reproductive technologies and global movements in the 21st century. Hum Reprod Update 2015; 21: 411–426. doi:10.1093/humupd/dmv016

[10] Niederberger C, Pellicer A, Cohen J et al. Forty years of IVF. Fertil Steril 2018; 110: 185–324. doi:10.1016/j.fertnstert.2018.06.005

[11] Bartnitzky S, Blumenauer V, Czeromin U et al. D·I·R Annual 2020 – The German IVF-Registry. J Reproduktionsmed Endokrinol 2021; 18: 203–247

[12] Zaninovic N, Rosenwaks Z. Artificial intelligence in human in vitro fertilization and embryology. Fertil Steril 2020; 114: 914–920. doi:10.1016/j.fertnstert.2020.09.157

[13] Grüber K, Gruisbourne BD, Pömsl J. Präimplantationsdiagnostik in Deutschland. Handreichung. Berlin: IMEW;2016 . Accessed December 23, 2021 at: https://www.imew.de/fileadmin/Dokumente/dokumente_2016/HandreichungZurPID_20160301.pdf

[14] Wang R, Pan W, Jin L et al. Artificial intelligence in reproductive medicine. Reproduction 2019; 158: R139–R154. doi:10.1530/REP-18-0523

[15] Raef B, Ferdousi R. A Review of Machine Learning Approaches in Assisted Reproductive Technologies. Acta Inform Med 2019; 27: 205–211. doi:10.5455/aim.2019.27.205-211

[16] Dengler K, Fangerau H (eds.). Zuteilungskriterien im Gesundheitswesen: Grenzen und Alternativen. Bielefeld: Transcript; 2013.

[17] Bernard A. Das Diktat der Fruchtbarkeit. Woopen C (ed.). Fortpflanzungsmedizin in Deutschland: Entwicklungen, Fragen, Kontroversen. Bonn: Bundeszentrale für politische Bildung; 2016: 11–19

[18] De Geyter C, Calhaz-Jorge C, Kupka MS et al. ART in Europe, 2015: results generated from European registries by ESHRE. Hum Reprod Open 2020; 2020: hoz038. doi:10.1093/hropen/hoz038

[19] Gleicher N, Kushnir VA, Albertini DF et al. Improvements in IVF in women of advanced age. J Endocrinol 2016; 230: F1–F6. doi:10.1530/JOE-16-0105

[20] Stoop D, Ermini B, Polyzos NP et al. Reproductive potential of a metaphase II oocyte retrieved after ovarian stimulation: an analysis of 23 354 ICSI cycles. Hum Reprod 2012; 27: 2030–2035. doi:10.1093/humrep/des131

[21] Noorbakhsh-Sabet N, Zand R, Zhang Y et al. Artificial Intelligence Transforms the Future of Health Care. Am J Med 2019; 132: 795–801. doi:10.1016/j.amjmed.2019.01.017

[22] Siristatidis C, Vogiatzi P, Pouliakis A et al. Predicting IVF outcome: A proposed web-based system using artificial intelligence. In Vivo 2016; 30: 507–512

[23] Mortimer ST, van der Horst G, Mortimer D. The future of computer-aided sperm analysis. Asian J Androl 2015; 17: 545–553. doi:10.4103/1008-682X.154312

[24] Gudeloglu A, Brahmbhatt JV, Parekattil SJ. Medical management of male infertility in the absence of a specific etiology. Semin Reprod Med 2014; 32: 313–318. doi:10.1055/s-0034-1375184

[25] Goodson SG, White S, Stevans AM et al. CASAnova: a multiclass support vector machine model for the classification of human sperm motility patterns. Biol Reprod 2017; 97: 698–708. doi:10.1093/biolre/iox120

[26] Girela JL, Gil D, Johnsson M et al. Semen parameters can be predicted from environmental factors and lifestyle using artificial intelligence methods. Biol Reprod 2013; 88: 99. doi:10.1095/biolreprod.112.104653

[27] Yanez L, Han J, Behr B et al. Human oocyte developmental potential is predicted by mechanical properties within hours after fertilization. Nat Commun 2016; 7: 10809. doi:10.1038/ncomms10809

[28] Cavalera F, Zanoni M, Merico V et al. A neural network-based identification of developmentally competent or incompetent mouse fully-grown oocytes. J Vis Exp 2018; 133: 56668. doi:10.3791/56668

[29] Saeedi P, Yee D, Au J et al. Automatic Identification of Human Blastocyst Components via Texture. IEEE Trans Biomed Eng 2017; 64: 2968–2978. doi:10.1109/TBME.2017.2759665

[30] Khosravi P, Kazemi E, Zhan Q et al. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. NPJ Digit Med 2019; 2: 21. doi:10.1038/s41746-019-0096-y

[31] Kaufmann SJ, Eastaugh JL, Snowden S et al. The application of neural networks in predicting the outcome of in vitro fertilization. Hum Reprod 1997; 12: 1454–1457. doi:10.1093/humrep/12.7.1454

[32] Durairaj M, Thamilselvan P. Applications of Artificial Neural Network for IVF Data Analysis and Prediction. JEC&AS 2013; 2: 11–15

[33] Song Z, Li W, O'leary S et al. Can the use of diagnostic and prognostic categorisation tailor the need for assisted reproductive technology in infertile couples? Aust N Z J Obstet Gynaecol 2021; 61: 297–303. doi:10.1111/ajo.13273

[34] Topol EJ. Welcoming New Guidelines for AI clinical research. Nat Med 2020; 26: 1318–1320. doi:10.1038/s41591-020-1042-x

[35] Nagendran M, Chen Y, Lovejoy CA et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ 2020; 368: m689. doi:10.1136/bmj.m689

[36] Genin K, Grote T. Randomized Controlled Trials in Medical AI. Philosophy of Medicine 2021; 2. doi:10.5195/philmed.2021.27

[37] Harper J, Magli MC, Lundin K et al. When and how should new technology be introduced into the IVF laboratory? Hum Reprod 2012; 27: 303–313. doi:10.1093/humrep/der414

[38] Kreuzer VK, Kimmel M, Schiffner J et al. Possible Reasons for Discontinuation of Therapy: an Analysis of 571 071 Treatment Cycles From the German IVF Registry. Geburtshilfe Frauenheilkd 2018; 78: 984–990. doi:10.1055/a-0715-2654

[39] Gameiro S, Boivin J, Peronace L et al. Why do patients discontinue fertility treatment? A systematic review of reasons and predictors of discontinuation in fertility treatment. Hum Reprod Update 2012; 18: 652–669. doi:10.1093/humupd/dms031

[40] Ergin RN, Polat A, Kars B et al. Social stigma and familial attitudes related to infertility. Turk J Obstet Gynecol 2018; 15: 46–49. doi:10.4274/tjod.04307

[41] Husain W, Imran M. Infertility as seen by the infertile couples from a collectivistic culture. J Community Psychol 2021; 49: 354–360. doi:10.1002/jcop.22463

[42] UN-Vollversammlung. Allgemeine Erklärung der Menschenrechte. Paris: Vereinte Nationen;10.12.1948 . Accessed October 12, 2021 at: http://www.un.org/en/universal-declaration-human-rights/

[43] Ethics Committee of the American Society for Reproductive Medicine. Disparities in access to effective treatment for infertility in the United States: an Ethics Committee opinion. Fertil Steril 2015; 104: 1104–1110. doi:10.1016/j.fertnstert.2015.07.1139

[44] Robertson JA. Children of Choice. Freedom and the New Reproductive Technologies. Princeton: Princeton University Press; 1994.

[45] Callahan S. The Ethical Challenges of the New Reproductive Technologies. Morrison EE (ed.). Health Care Ethics: Critical Issues for the 21 st Century. 2 ed. Sudbury, MA: Jones and Bartlett Publishers; 2009: 71–86

[46] Coiera E. On algorithms, machines, and medicine. Lancet Oncol 2019; 20: 166–167. doi:10.1016/S1470-2045(18)30835-0

[47] Rockliff HE, Lightman SL, Rhidian E et al. A systematic review of psychosocial factors associated with emotional adjustment in in vitro fertilization patients. Hum Reprod Update 2014; 20: 594–613. doi:10.1093/humupd/dmu010

[48] Deutsches IVF-Register. Jahrbuch 2019. J Reproduktionsmed Endokrinol 2020; 5: 199–239. Accessed December 23, 2021 at: https://www.deutsches-ivf-register.de/perch/resources/dir-jahrbuch-2019-de.pdf

[49] Ahuja AS. The Impact of Artificial Intelligence in Medicine on the Future Role of the Physician. PeerJ 2019; 7: e7702. doi:10.7717/peerj.7702

[50] de Miguel I, Sanz B, Lazcoz G. Machine learning in the EU health care context: exploring the ethical, legal and social issues. iCS 2020; 23: 1139–1153. doi:10.1080/1369118X.2020.1719185

[51] Cohen IG, Amarasingham R, Shah A et al. The legal and ethical concerns that arise from using complex predictive analytics in health care. Health Aff (Millwood) 2014; 33: 1139–1147. doi:10.1377/hlthaff.2014.0048

[52] Ranisch R. Consultation with Doctor Twitter: Consent Fatigue, and the Role of Developers in Digital Medical Ethics. Am J Bioeth 2021; 21: 24–25. doi:10.1080/15265161.2021.1926595

[53] Kohli M, Prevedello LM, Filice RW et al. Implementing machine learning in radiology practice and research. AJR Am J Roentgenol 2017; 208: 754–760. doi:10.2214/AJR.16.17224

[54] Price WN. Big Data and Black-Box Medical Algorithms. Sci Transl Med 2018; 10. doi:10.1126/scitranslmed.aao5333

[55] Rudin C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. Nat Mach Intell 2019; 1: 206–215. doi:10.1038/s42256-019-0048-x

[56] Gaube S, Suresh H, Raue M et al. Do as AI say: susceptibility in deployment of clinical decision-aids. NPJ Digit Med 2021; 4: 31. doi:10.1038/s41746-021-00385-9

[57] Calhaz-Jorge C, De Geyter CH, Kupka MS et al. Survey on ART and IUI: legislation, regulation, funding and registries in European countries: The European IVF-monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE). Hum Reprod Open 2020; 2020: hoz044. doi:10.1093/hropen/hoz044

[58] Passet-Wittig J, Bujard M. Medically assisted Reproduction in developed Countries: Overview and societal Challenges. Schneider NF, Kreyenfeld M (eds.). Research Handbook on the Sociology of the Family. Cheltenham: Edward Elgar Publishing; 2021: 417–438. doi:10.4337/9781788975544

[59] Rauprich O, Berns E, Vollmann J. Who should pay for assisted reproductive techniques? Answers from patients, professionals and the general public in Germany. Hum Reprod 2010; 25: 1225–1233. doi:10.1093/humrep/deq056

[60] Maddox TM, Rumsfeld JS, Payne PRO. Questions for Artificial Intelligence in Health Care. JAMA 2019; 321: 31–32. doi:10.1001/jama.2018.18932

[61] Ross L. Reproductive Justice: An Introduction. Oakland: University of California Press; 2017.

[62] Mau S. Das metrische Wir. Über die Quantifizierung des Sozialen. Frankfurt a.M.: Suhrkamp; 2017.

[63] Obermeyer Z, Powers B, Vogeli C et al. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019; 366: 447–453. doi:10.1126/science.aax2342

[64] Barocas S, Selbst AD. Big data's disparate impact. Calif Law Rev 2016; 104: 671. doi:10.2139/ssrn.2477899

[65] Castelvecchi D. Can we open the black box of AI? Nature 2016; 538: 20–23. doi:10.1038/538020a

[66] Cohen IG. Informed Consent and Medical Artificial Intelligence: What to Tell the Patient? Georgetown Law J 2020; 108: 1425–1469. doi:10.2139/ssrn.3529576

[67] Wöhlke S, Schicktanz S. 'Special Issue: Why Ethically Reflect on Empirical Studies in Empirical Ethics? Case Studies and Commentaries'. J Empir Res Hum Res Ethics 2019; 14: 424–427. doi:10.1177/1556264619862395

[68] Turley P, Meyer MN, Wang N et al. Problems with Using Polygenic Scores to Select Embryos. N Engl J Med 2021; 385: 78–86. doi:10.1056/NEJMsr2105065

[69] Heintz B. Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs. Z Soziol 2010; 39: 162–181. doi:10.1515/zfsoz-2010-0301

[70] Boyd PA, Devigan C, Khoshnood B et al. Survey of prenatal screening policies in Europe for structural malformations and chromosome anomalies, and their impact on detection and termination rates for neural tube defects and Down's syndrome. BJOG 2008; 115: 689–696. doi:10.1111/j.1471-0528.2008.01700.x

[71] Ranisch R. Liberale Eugenik? Kritik der selektiven Reproduktion. Berlin: Metzler; 2021.