

Towards deep phenotyping pregnancy: a systematic review on artificial intelligence and machine learning methods to improve pregnancy outcomes

Lena Davidson and Mary Regina Boland

Corresponding author: Mary Regina Boland, University of Pennsylvania, 423 Guardian Drive, 421 Blockley Hall, Philadelphia, PA 19104, USA.
Tel.: +1 215-5737394; E-mail: bolandm@upenn.edu

Abstract

Objective: Development of novel informatics methods focused on improving pregnancy outcomes remains an active area of research. The purpose of this study is to systematically review the ways that artificial intelligence (AI) and machine learning (ML), including deep learning (DL), methodologies can inform patient care during pregnancy and improve outcomes.

Materials and methods: We searched English articles on EMBASE, PubMed and SCOPUS. Search terms included ML, AI, pregnancy and informatics. We included research articles and book chapters, excluding conference papers, editorials and notes.

Results: We identified 127 distinct studies from our queries that were relevant to our topic and included in the review. We found that supervised learning methods were more popular ($n = 69$) than unsupervised methods ($n = 9$). Popular methods included support vector machines ($n = 30$), artificial neural networks ($n = 22$), regression analysis ($n = 17$) and random forests ($n = 16$). Methods such as DL are beginning to gain traction ($n = 13$). Common areas within the pregnancy domain where AI and ML methods were used the most include prenatal care (e.g. fetal anomalies, placental functioning) ($n = 73$); perinatal care, birth and delivery ($n = 20$); and preterm birth ($n = 13$). Efforts to translate AI into clinical care include clinical decision support systems ($n = 24$) and mobile health applications ($n = 9$).

Conclusions: Overall, we found that ML and AI methods are being employed to optimize pregnancy outcomes, including modern DL methods ($n = 13$). Future research should focus on less-studied pregnancy domain areas, including postnatal and postpartum care ($n = 2$). Also, more work on clinical adoption of AI methods and the ethical implications of such adoption is needed.

Key words: literature review; pregnancy; artificial intelligence; machine learning

Introduction

In the field of medicine, the theory of ‘joint decision-making’ between humans and artificial intelligence (AI) holds the

promise of promoting high-performance medicine. AI is a broad science of simulating the natural intelligence demonstrated by humans and animals via artificial means. Machine learning (ML)

Lena Davidson received her MS degree at College of St. Scholastica, Duluth, MN, USA. Presently, she is a data analyst in the Department of Biostatistics, Epidemiology, and Informatics at the University of Pennsylvania.

Mary Regina Boland is an assistant professor of informatics in the Department of Biostatistics, Epidemiology, and Informatics at the University of Pennsylvania. He specializes in developing informatics methods that utilize data from Electronic Health Records to reveal novel clinical and scientific findings. Principally, these methods involve novel machine learning, artificial intelligence and language processing techniques.

Submitted: 22 May 2020; **Received (in revised form):** 13 October 2020

© The Author(s) 2021. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

For commercial re-use, please contact journals.permissions@oup.com

is a subset of AI, in the field of computer science. ML refers to a number of methods and algorithms, and different learning types: supervised, unsupervised and reinforcement learning [1, 2]. In the early days of AI in medicine, AI systems were standalone systems without direct connection to electronic health records (EHRs). Today, medicine is progressing towards a learning health system in which knowledge derived from information EHR data can be directly applied to care. Digitized clinical data in the EHR, genomics and biology present a wealth of information, open new opportunities and come with new challenges. AI and informatics methodologies are critical to enabling the learning health system.

One of the ways that AI and ML can be used in healthcare is in enabling 'deep phenotyping'. In medicine, a disease phenotype refers to a deviation from healthy morphology or physiology [3]. Study of phenotypes requires knowledge of the spectrum of phenotypes associated with a disease entity. Defining what constitutes a diseased phenotype versus a healthy phenotype is often challenging. Deep phenotyping disease is a step towards precision medicine in which a comprehensive and precise phenotyping of disease presentation takes place. The individual components of the phenotype are observed, described and analyzed in order to develop knowledge of human disease. AI and ML methodologies naturally apply to characterize phenotypes in a 'deep' manner including multiple methodologies (e.g. genetics, imaging, diagnostics and so forth). These methods can exhaustively examine data with high granularity and dimensionality and make use of the broad range of data types that may be processed in deep phenotyping. For instance, nuanced phenotypic traits may be more readily available in unstructured data (e.g. clinical notes), requiring natural language processing (NLP) to identify relative information [4]. Several diseases and adverse outcomes during pregnancy (e.g. preterm birth, preeclampsia and miscarriage) have complicated and difficult to understand etiology, leaving little to be done for prevention [5–7]. Deep phenotyping these patient states (i.e. pregnancy phenotypes) could help improve adverse outcomes and provide further insight into diseases during pregnancy.

AI and ML methods can be employed to enable deep phenotyping, especially with regards to the pregnancy state where many different data types are used (e.g. ultrasound imaging, diagnostic screening, fetal monitoring, genetics). AI and ML methods in medicine are an emerging field and have been described in detail of its theory and current applications across several medical disciplines [8, 9] and many disease areas and clinical states [8, 10]. However, the use of AI to improve women's health, specifically during the pregnancy, has had limited clinical use. In the 2019 International Medical Informatics Association Yearbook of Medical Informatics, there are no research articles focused on pregnancy and maternal health, illustrating the lack of research focus on this important aspect of women's health [11]. For researchers and clinicians alike, AI techniques have promise to derive sound results and improve care at each stage of pregnancy [12–14].

Overall, the purpose of this study is to systematically review the ways that AI and ML methodologies can inform patient care during pregnancy and improve outcomes. We seek to (a) describe which medical fields and informatics areas apply AI and ML, (b) find where in pregnancy are these methodologies used, (c) describe clinical decision support systems (CDSSs) that employ AI or ML and (d) identify literature gaps for future research.

Methods

We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines when conducting our literature review [15]. See Supplementary Table available online at <https://academic.oup.com/bib> for PRISMA 2009 Checklist.

Systematic review of literature

We searched three databases: *EMBASE*, *PubMed* and *SCOPUS*. *PubMed* is a freely available database housing publications maintained by the United States National Center for Biotechnology Information. *PubMed* facilitates searching across three National Library of Medicine resources; the largest component is *MEDLINE*, followed by content from *PubMed Central (PMC)* and *Bookshelf* [16]. *EMBASE* is a bibliographic database focused on pharmacovigilance. *SCOPUS* is Elsevier's publication database containing articles from over 36 000 peer-reviewed journals. We used site licenses from the University of Pennsylvania libraries to search *SCOPUS* and *EMBASE*. The initial query was broad: in *PubMed*, we searched the keywords in all fields; in *SCOPUS*, we searched article title, abstract and keywords; and in *EMBASE*, we used quick search. On 5 December 2019, we used the following search query to cast a wide net of potentially relevant papers for inclusion of our review: AI AND pregnancy. On 18 February 2020, we completed a more focused query, in order to identify papers that focused on ML: informatics AND pregnancy AND AI AND ML.

The *PubMed* interface uses Automatic Term Mapping to automatically map search words to their respective Medical Subject Headings terms [17]. After retrieving results from each database, we removed duplicate studies using exact *PubMed* ID match. When *PubMed* IDs were absent, we reviewed articles, comparing title, author list and publication date to further identify duplicate publications. We filtered the results by excluding non-English studies, conference papers, editorials and notes.

The eligibility criteria include AI and maternal health; papers were categorized by pregnancy stage and health concern. No location or publication date restrictions were applied and no unpublished papers were retained. Retrieved articles were inspected by an independent review (L.D.), and in cases where the eligibility of the study was unclear, it was reviewed by a second reviewer (M.R.B.). We excluded studies for which we were not able to gain access to the manuscripts.

Results

Systematic literature review results

First, we searched *EMBASE*, *PubMed* and *SCOPUS* for articles on pregnancy and AI. We found 245 from *SCOPUS*, 181 from *EMBASE* and 128 from *PubMed*. We removed duplicate studies. In total, we found that 381 distinct research papers were retrieved on the first query. For the second query, we found additional papers, including 4 from *SCOPUS*, 4 from *EMBASE* and 46 from *PubMed*. After removing duplicates from the new query, we had a result set of 427 distinct papers. Of these 427 distinct research studies, we excluded the 9 non-English studies, 124 conference papers, 4 editorials and 5 notes, resulting in a set of 285 research papers (Figure 1).

The next step was to assess the 285 remaining research articles for relevance. We manually reviewed the 285 articles to determine if they met selection criteria: (1) focused on AI, (2) related to pregnancy and (3) related to healthcare and health

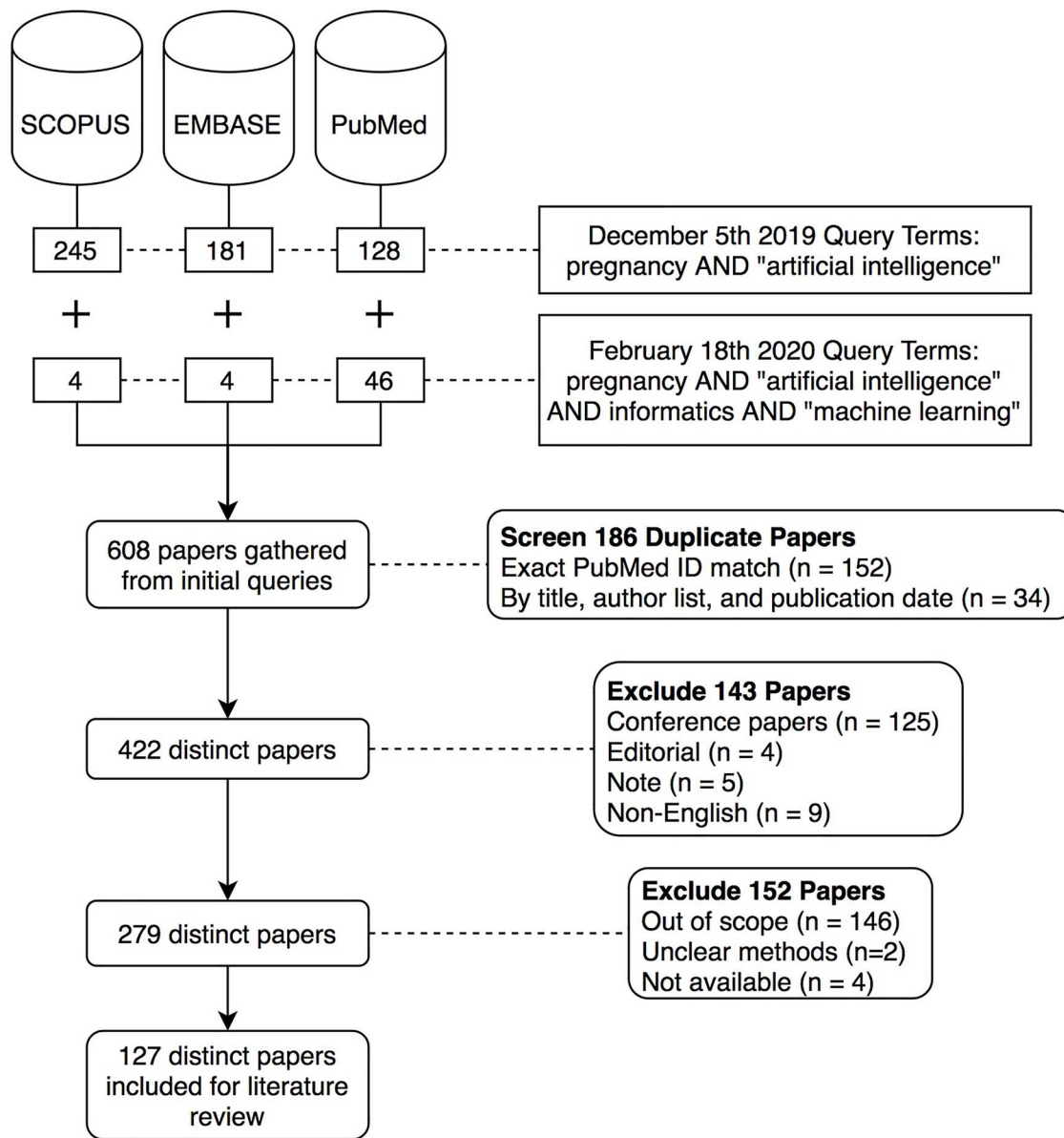


Figure 1. PRISMA diagram showing our literature search inclusion process.

information. These inclusion criteria resulted in 127 relevant papers. The final set contains almost four decades of research, issued from 1982 to 2020.

Identification of themes

We grouped the studies by the stage in pregnancy: preconception, assisted reproductive technology (ART), prenatal screening and monitoring, preterm labor, birth and delivery (full term) and postnatal to illustrate the different clinical areas where AI and ML methods were applied. These areas are shown in Figure 2; we distinguish between pregnant person and embryo, fetus or neonate as methods are designed explicitly for each respective concern. References to maternal and women's health are reflecting the terms and concepts of research included in this paper. Pregnant people include transgender and gender non-conforming individuals; however, no papers ($n=0$ studies) define this population explicitly.

CDSSs are explicitly described in each category where they are present. Table 1 presents studies by pregnancy stage, with an overview of the methodology and results. Not included in Table 1 is one evaluation of an application described in an included paper [18], and 8 review papers, reviewing applications in the following: general obstetrics [19], translational science [20], intrapartum surveillance [21, 22], fetal heart rate (FHR) [23], labor fetal assessment [24], ART [12] and mobile health (mHealth) for antepartum care [25]. AI may help address challenges in birth defects research. [20]. Reviews found great promise in AI applications, but not realized in clinical care [21, 22, 25], or missing external validation [12]. AI applications are warned as adjunct, not replacement, of healthcare professionals [19, 23]. One review found no evidence supporting improved pregnancy outcomes with AI applications [24]. AI methods applied are shown in Figure 3. Refer to Table 2 for further references into the methodologies observed in this review, along with pregnancy and informatics domains in which these methods were applied. Of the 127 studies, 16 applied AI and ML methods to retrospective

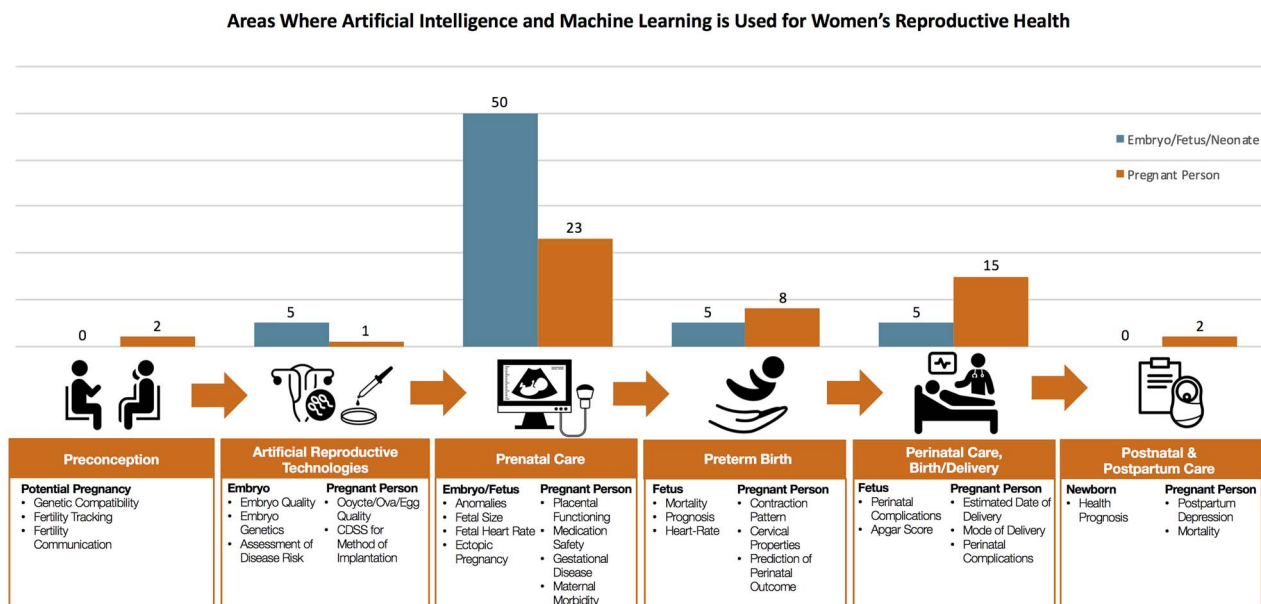


Figure 2. Different stages along the pregnancy timeline where AI and ML methods are used. This figure illustrates the points of care where AI and ML have been applied at each stage of pregnancy: preconception; artificial reproductive technologies; prenatal care; preterm birth; perinatal care, birth and delivery; and postnatal care. Healthcare concerning the pregnant person and fetus is distinguished for ease of reading. We also include the distribution of studies by pregnancy stage to illustrate the stages of the pregnancy with the most number of studies for easy identification of literature gaps.

EHR data. A total of 12 studies used expert systems as their AI methods (Figure 3), and these are often older, more historical AI methods applied to women's health. Studies tended to focus on the prenatal care aspect of the pregnancy (Figure 2), with more studies focused on fetal aspects of the prenatal care ($n = 50$ studies) versus the maternal aspects of care ($n = 23$ studies). Another popular area where AI methods were used was perinatal/birth and delivery (15 studies focused on maternal outcomes and 5 focused on fetal outcomes). Other stages are represented to a lesser extent, especially preconception care and postpartum care with two studies each (Figure 2).

Pregnancy care spans across several disciplines, some of which overlap: electronic monitoring, radiology and imaging, mHealth, CDSS, NLP and search analytics, genetic and chromosomal abnormalities, pregnancy complications, labor and delivery, and postpartum care. Electronic monitoring was the most common with 27 studies (Figure 4). Radiology and CDSS were also popular areas for using AI methods. Informatics areas that are only beginning to utilize AI methods for pregnancy-related care include mHealth ($n = 9$), chromosomal anomaly detection ($n = 7$) and NLP ($n = 3$).

Discussion

Mobile health

Development of provider-facing and direct-to-consumer mHealth applications supports preconception [26], gestational disease management [27], remote health monitoring [28], low-resource prenatal care [29–31], text messaging [32–34], patient education [32, 35], fetal health status prediction [36], preeclampsia prediction [37] and perinatal depression [34]. A 2019 systematic review determined mHealth applications with promising potential for use by healthcare workers during antenatal care [25]: Babyscripts [32], OpenSRP [29], PANDA [28], PotM [37], mHealth Guatemala

[30], Expect With Me [35], mPAMANECH [31] and COMM CARE [38]. A CDSS with mobile patient support predicts normal, pathological and potentially pathological fetal health status; significant features found include fetal age, maternal age, blood serotype, delivery number and illnesses regarding the current pregnancy [36]. Together, boosted decision tree (DT), decision forest (DF) and decision jungle were found to be the most efficient.

Direct-to-consumer mHealth applications have potential to increase engagement and empower pregnant people in their healthcare. Three studies arose, focusing on predicting the fertility window [26], text mining to understand communications with a sexual and reproductive health information service [33] and a mHealth text messaging system for perinatal depression [34]. It should be noted that in these studies the authors declare financial conflict of interest. A text mining approach was developed using naïve Bayes (NB) and basic NLP techniques, to understand how Kenyan men and women communicated with askNivi, a free sexual and reproductive health information service [33]. The users wrote most often about family planning methods, contraception, side effects, pregnancy, menstruation and sex. A majority of users sought factual information, followed by requests for advice and reporting symptoms. A prototype mHealth text messaging system for perinatal depression, Tess, was tested with mothers recruited from public hospitals outside of Nairobi, Kenya [34]. Prediction of a woman's fertile window through data received by a wearable bracelet achieved 90% accuracy using a random forest (RF) classifier [26]. A prospective longitudinal study determined what phase-based differences a wearable bracelet could detect in users' wrist skin temperature, heart rate, heart rate variability, respiratory rate and skin perfusion [26].

Applications address the needs of the pregnant person, support clinical care and often aim to limit in-person visits without compromising quality of care. Post the COVID-19 lockdown, development and application of mHealth and telehealth for

Table 1. Articles published using AI to improve pregnancy outcomes: method details, pregnancy area, health issue, dataset details and model results

Study	Methods	Objective	Health issue	Dataset	Result
Preconception Goodale et al. (2019)	Residual maximum likelihood, Satterthwaite degrees of freedom, Bonferroni correction, RF	To develop an algorithm that could identify the fertile window in real time	Fertility window	237 conception-seeking Swiss women	Can detect significant, concurrent phase-based shifts in WST, heart rate and respiratory rate (all $P < 0.001$), while HRV and skin perfusion show less significance ($P < 0.05$)
Green et al. (2019a)	NB, NLP	Apply a text-mining approach to characterize the ways that Kenyan men and women communicated with the first iterations of askNivi, a free sexual and reproductive health information service	Reproductive health	179 609 total messages (87 180 inbound and 92 429 outbound) from askNavi database	Users wrote most often about family planning methods, contraception, side effects, pregnancy, menstruation and sex. Most users sought factual information, but requests for advice and symptom reports were common
ART Morales et al. (2008)	Bayes' classifiers: NB, selective NB, semi NB, tree-augmented NB, k-dependence Bayesian classifier	Proposes the application of Bayesian classifiers to the embryo selection problem	IVF outcome	63 clinical files from IVF program in Clinica del Pilar in San Sebastián, (July 2003–December 2005, aged 27–46) and morphological characteristics of embryos	General accuracy of 68.25% with 95.55% of specificity and no-implantation predictive value of 70.49%
Nyalwidhe et al. (2013)	Genetic algorithm, SVM, k-NN	To identify differentially expressed peptides/proteins in the culture media of embryos grown during IVF treatment to establish their value as biomarkers predictive of implantation potential and live birth	Pregnancy success	Micro-droplets of embryo culture media from IVF patients (conditioned) and control media (January 2006–December 2010)	Found downregulation of apolipoprotein A-1 peptides and apolipoprotein A-1 protein in the pregnancy groups
Uyar et al. (2015)	NB, DT, k-NN, SVM, MLP, RBF	To predict implantation outcome of individual embryos in an IVF cycle with the aim of providing decision support on the number of embryos transferred	Pre-implantation	2453 embryos transferred at day 2 or day 3 after intracytoplasmic sperm injection (ICSI)	NB, accuracy = 80.4%, sensitivity = 63.7%, false alarm rate = 17.6%
Milewski et al. (2017)	Principal component analysis, ANN	To answer the question of how much information about embryo implantation potential can be obtained from morphokinetic parameters	Pre-implantation	Time-lapse recordings of 610 embryos	Validation set, AUC = 0.71
Khosravi et al. (2019)	DNN, Google's Inception Model, chi-squared automatic interaction detection (CHAID) DT	Implemented an AI approach to select highest quality embryos using a large collection of human embryo time-lapse images	Pre-implantation	Time-lapse images from 10 148 embryos from the Center for Reproductive Medicine at Weill Cornell Medicine (2012/05–2017/12)	Chance of pregnancy based on individual embryos varies from 13.8% (age ≥ 41 and poor quality) to 66.3% (age < 37 and good quality)

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Tran et al. (2019)	DL	To investigate if a DL model named IVY is a valid tool for the prediction of the implantation potential of human pre-implantation embryos	Pre-implantation	Time-lapse videos and clinical outcomes of 10 638 embryos from eight different IVF clinics, across four different countries (January 2014–December 2018)	5-fold cross validation, AUC = 0.93 hold-out validation test, AUC ranges from 0.90 to 0.95
Prenatal Bassil and Dripps (1989)	Interpolation at the comb filter output stage, peak selection algorithm	Facilitate safe long-term patient monitoring and reduce erroneous estimates during periods of now signal-to-noise ratio	Fetal heart rate	Phonocardiographic signals	N/A
Hernando et al. (1996)	Causal probabilistic network	Describes the integration of qualitative and quantitative reasoning modules within the DIABNET advisory system	CDSS	9 GDM patients, blood glucose samples—52 cases for review	82.8% of cases correctly identified
Nagy et al. (1997)	Rule-based (static and dynamic knowledge base)	To automatically interpret serological test results of toxoplasma infection	Fetal health	Retrospective study group of 1000 pregnant women	Latent infections correctly classified = 99.10%
Alonso-Betanzos et al. (1999)	Bayes' model, discriminant analysis, ANNs and the Shortliffe and Buchanan uncertainty-based model	Compare developed NST-EXPERT system against other ML methods in prediction of fetal outcome	Fetal health	3386 cases, including fetal-maternal risk features and early neonatal status parameters	Proposed model is less flexible than ANN method but does not depend on expert input
Meinecke et al. (2002)	ICA, BSS	Propose resampling methods to tackle this question and illustrate their usefulness for blind-source separation (BSS)	Fetal health	2500 points sampled at 500 Hz with eight electrodes located on abdomen and thorax of a pregnant woman	JADE is able to separate the heartbeat of the fetus from the heartbeat of the mother
Polak et al. (2004)	ANN	To develop novel GDM diagnosis system based on ANN	Maternal diagnosis	Epidemiological database obtained from National Center for Health Statistics (NCHS), Center for Disease Control and Prevention (CDC)	70% = true positive diagnoses
Kagan et al. (2008)	Likelihood ratio, multiples of the median, regression analysis	To examine the performance of first-trimester screening for trisomies 21, 18 and 13	Chromosomal abnormality	Singleton pregnancies (at 11 ⁺⁰ -13 ⁺⁶ weeks), including 56 376 normal cases, 395 with trisomy 21, 122 with trisomy 18 and 61 with trisomy 13 300 pregnant women	Overall FPR was 3.1% and the DRs of trisomies 21, 18 and 13 were 91, 97 and 94%, respectively
Etikan and Çağlar (2005)	Stepwise linear regression model, neural network with MLPs	To determine more accurate prediction method between linear and non-linear methods for prediction of babies' birth weight among maternal demographic characteristics	Birth weight		LR R ² = 59.8% Non-LR R ² = 59.8%

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Mantini et al. (2005)	ICA	To describe an automatic method that restores the fetal cardiac trace from fMCG recordings	Fetal heart function	Multichannel fMCG datasets of 66 healthy and 4 arrhythmic fetuses	Validation performed on healthy fetuses
Assaleh (2007)	Fuzzy logic, neural network, ANFIS	Investigate the use of ANFIS for fetal electrocardiogram extraction from two ECG signals recorded at the thoracic and abdominal areas of the mother's skin	Fetal health	Synthetic ECG data and 1 case of real ECG data	Performed well to suppress maternal components from fetal signal
Comani et al. (2007)	ICA, approximate entropy (ApEn) and sample entropy (SampEn)	Discusses a novel automated system based on entropy estimators for the classification of independent components (ICs)	Fetal heart function	40 fMCG datasets of normal fetuses with the gestational age ranging from 22 to 37 weeks	ApEn = detection rate of 98.75% SampEn = detection rate of 97.92%
Karvounis et al. (2007)	Time-frequency (T-F) analysis, complex wavelets and matching theory techniques, heuristic algorithm, histogram-based technique	Introduces an automated methodology for the extraction of fetal heart rate (FHR) from cutaneous potential abdominal electrocardiogram (abdECG) recordings	Fetal heart function	8 short and 10 long-duration abdECG recordings, obtained between the 20th and the 41st week of gestation	Accuracy = 97.47%
Singh et al. (2007)	Data management resources, comparative bioinformatics	Conceptualize a system that can help address the new challenges posed by the translational science of birth defects research	Birth defects	Mouse model	N/A
Mosaliganti et al. (2007)	Surface pocket modeling	Propose a technique for detecting pockets on a surface-of-interest	Placenta functioning	Case study of genetically modified mouse placenta	Sensitivity = 0.84 Specificity = 0.90
Yu et al. (2009)	Support vector regression (SVR), fuzzy logic, non-dominated sorting genetic algorithm (NSGA)	A novel method based on the SVR is proposed to improve the weight estimation accuracy for fetuses of less than 2500 g	Fetal weight and low birth weight	Records of 386 fetuses with ABW less than 2500 g (Shanghai First Maternity and Infant Health Hospital, China)	Proposed model produces significant improvement (1.9%–4.2%) on the accuracy of fetal weight estimation over several widely used formulas
Evans and Cuckle (2011)	Multiples of the gestation-specific median (MoMs)	Develop a method to allow for differences in quality among ultrasound nuchal translucency providers or laboratories, in order to improve performance	Chromosomal abnormality	7237 women (median age of 36 at delivery), who had first trimester screening, including 32 cases of Down syndrome	Performance demonstrated in 5 theoretical examples
Marttala et al. (2011)	Multiples of medians	To examine the performance of first-trimester combined screening after adding the specific algorithms for trisomies 18 and 13 in the Down syndrome screening program for chromosomal abnormalities other than trisomy 21 and to determine the outcomes of such pregnancies	Chromosomal abnormality	Pregnant women (n = 56 076), first-trimester combined Down syndrome screening in Northern and Eastern Finland (1 June 2002–31 December 2008)	Adding a T13 algorithm DR improved from 36.4 to 54.5%, with an increase of FPR from 4.3 to 4.5%. T18 algorithm improved the DR from 55.6 to 74% with the same increase of FPR of 0.2%

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Tejera et al. (2011)	ANN	A model construction for classification of women with normal, hypertensive and preeclamptic pregnancy in different gestational ages using maternal heart rate variability (HRV) indexes	Hypertensive disorders	568 ECG records	Sensitivity for preeclampsia around 80%, specificity = 85–90%
Boisvert et al. (2012)	Bayesian classification	To identify, using capillary electrophoresis and chemometrics, early biomarkers in human amniotic fluid of large-for-gestational-age (LGA) infants	Birth weight	109 mother–infant pairs, pregnant women undergoing routine age-related amniocentesis between 12 and 20 gestational weeks at St Mary's Hospital Center in Montreal, QC, Canada	LGA ($n = 23$) was differentiated from appropriate for GA ($n = 86$) with a sensitivity of 100% and a specificity of 98%
Gebeh et al. (2012)	geNorm, Normfinder	To evaluate the stability of 12 reference genes and suggest those that are stable for use as internal control genes in fallopian tubes and endometrium from ectopic pregnancy and healthy non-pregnant controls	Ectopic pregnancy	Fallopian tubes and endometrial biopsies were obtained from a total of 12 volunteers	Ranked top 6 by geNorm and Normfinder: UBC, GAPDH, CYC1 and EIF4A2 (fallopian tubes) and UBC and ATP5B (endometrium)
Ocak (2013)	SVM, genetic algorithm	To evaluate fetal well-being from the CTG recordings using ML methods	Fetal health	1831 CTG recordings from the UCI Machine Learning Repository (1655 normal fetal state, 176 pathological state)	Predict fetal state as normal, accuracy = 99.3%, or pathological, accuracy = 100%
Yeo and Romero (2013)	Spatiotemporal image correlation, 'intelligent navigation' technology	To explore visualization of standard fetal echocardiography views from volume datasets obtained with spatiotemporal image correlation (STIC)	Fetal health	Tested on 50 STIC volume datasets, echocardiography	A combination of diagnostic planes and/or VIS-Assistance® in 98–100% in normal cases
Haghanahi and Borkholder (2014)	QRS annotation, Gaussian model selection, Kalman filtering	Describe an algorithm to automatically locate QRS complexes in non-invasive fetal ECG signals	Fetal health	Obtained from a set of four electrodes placed on the mother's abdomen, ECG datasets provided by the PhysioNet/CinC challenge 2013	SE = 98.22% and PPV = 97.75%
Varanini et al. (2014)	ICA, weighted singular value decomposition	Present an efficient unsupervised algorithm for fetal QRS complex detection from abdominal multichannel signal recordings combining ICA and maternal ECG cancelling	Fetal health	447 records extracted from five abdominal signal collections acquired, fECG datasets provided by the PhysioNet/CinC challenge 2013	Subset of 69 records, SE = 99.4% and PPV = 99.2%
Sanderson et al. (2014)	Virtual organ computer-aided analysis (VOCAL), sub-noise gain (SNG)	Asses the intra-observer and inter-observer repeatability of selecting the SNG level when acquiring placental volumes with 3D power Doppler for analysis using VOCAL	Tissue perfusion	240 static 3D volumes from 60 women with uncomplicated singleton pregnancies (20–38 weeks of gestation)	SNG exhibited excellent intra-observer and inter-observer reliability
Chu et al. (2014)	Non-invasive prenatal tests, whole genome sequencing	To explore the lower limits of microdeletion size detection via non-invasive prenatal tests using minimally invasive karyotyping (MINK) and introduce/evaluate the GC Content Random Effect Model (GCREM)	Genetic diagnosis	Maternal plasma DNA, amniocentesis at 21 weeks of gestation	The combination of targeted capture and the GCREM algorithm makes it possible to detect relatively small fetal mutations with low sequencing cost

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Kutarnia and Pederson (2015)	Markov random field (MRF)	Develop an algorithm to produce anatomically correct abdominal ultrasound volumes for use in obstetrics ultrasound simulators	Fetal health	FAST/ER FAN dataset Visible Human Female CT dataset	A fusion approach to MRF registration can produce accurate displacement fields much faster than standard approaches
Re et al. (2015)	LR, VOCAL	To evaluate the potential of VOCAL to estimate fetal thymus volume in normal pregnancies	Fetal health	3D volume data sets from the axial upper mediastinal section from 37 normal pregnancies (12–35 weeks' gestation) 443 placental images	Identification of the borders of the thymus and calculation of its volume were successful in 28 patients (77.7%) AUC = 96.77%, sensitivity = 98.04%, specificity = 93.75%
Lei et al. (2015)	Gaussian mixture model (GMM), Fisher vector (FV)	Propose a method to automatically stage placenta maturity from B-mode ultrasound images based on dense sampling and novel feature descriptors	Placental maturity		
Nambuete et al. (2015)	Regression forest, surface parametrization	Propose an automated framework for predicting gestational age (GA) and neurodevelopmental maturation of a fetus based on 3D ultrasound brain image appearance	Gestational age	Two study databases: INTERGROWTH-21st and INTERBIO-21st	Age prediction results strongly correlate with true GA confirming the link between maturational progression and neurosonographic activity observable across gestation
Gentillon et al. (2016)	Wavelet-based parameter, histogram-based parameters, Fisher texture analysis in three-dimensional space	To objectively compare the quality of 1.5tesla (1.5T) versus 3T MRI of the developing human brain (Figure 1), in order to determine whether the extra administrative cost is worthy for the patient and the healthcare system	Fetal health	288 MR images of normal fetal brain were used in the main assessment and 72 MR images of normal fetal brain (gestational age 20–40 weeks)	Both encoding modes are possible with both units, but higher 3T resolution is the main difference
Neocleous et al. (2016)	ANN, SVM, k-NN	To examine the potential value of using ML for the non-invasive estimation, at 11–13 weeks of gestation, the risk for euploidy, trisomy 21 (T21) and OCA, from suitable sonographic, biochemical markers and other relevant data	Chromosomal abnormalities	A database of 51 208 singleton pregnancy cases, while undergoing first trimester screening for aneuploidies	ANN system correctly identified all T21 cases 0% false-negative rate (FNR) and 96.1% of euploidies, i.e. 3.9% false-positive rate (FPR)
Boland et al. (2017)	ICD-9, ATC, LR, RF, multi-dimensional scaling	To classify these drugs into harmful and safe categories using knowledge gained from chemoinformatics and empirical data	Medication safety	Fetal loss cohort: 14 922 affected and 33 043 unaffected pregnancies Congenital anomalies cohort: 5658 affected and 31 240 unaffected infants	Out-of-bag accuracy of 91% for fetal loss and 87% for congenital anomalies
Neocleous et al. (2017)	ANN	To introduce a non-invasive diagnosis procedure for aneuploidy and to minimize the social and financial cost of prenatal diagnosis tests that are performed for fetal aneuploidies in an early stage of pregnancy	Chromosomal abnormalities	Total of 122 362 euploid and 967 aneuploid cases	100% detection rate for Trisomy 21 80% DR for Trisomies 13 and 18, Turner and Triploid syndromes
Pangalos et al. (2016)	Whole exome sequencing (WES)	Report the development and first applications of an expanded exome sequencing-based test, coupled to a bioinformatics-driven prioritization algorithm, targeting gene disorders presenting with abnormal prenatal ultrasound findings	Genetic abnormalities	14 euploid fetuses, from 11 on-going pregnancies and 3 products of abortion	Definitive or highly likely diagnosis was made in 6 of 14 cases (43%)

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Abassi et al. (2017)	'fingerprint of uncertainty' (FOU), type-2 fuzzy logic system (FLS) classifier, wavelet transform (WT)	To detect sharp wave transients that exist in the latent phase after hypoxia-ischemia (HI) in the electroencephalogram (EEG) in order to determine if these microscale transients are potential biomarkers of HI	Fetal health	EEG recordings of preterm fetal sheep model cohort	Best performing Type-2 FLS classifier occurred when the FOU's were built from Gaussian 2, scale 32 wavelet transformations of the sharp waves for a 1024 Hz sampled EEG.
Luo et al. (2017)	Weight SVM, weighted RF, LR	To develop and validate ML models for predicting, before and during pregnancy, women's risks of bearing children with CHD	Fetal health	Dataset of 33 831 live births, which contained 78 cases of CHD	TPR and TNR of the three classifiers were above 0.65 and 0.93
Fiorentino et al. (2017)	Genome-wide cfDNA analysis	Expanded conventional cfDNA-based non-invasive prenatal testing (NIPT) to cover the entire genome	Chromosomal abnormalities	12 114 pregnant women (>10 weeks of gestation) undergoing NIPT for common fetal aneuploidy, whole-blood samples (10 ml)	Detection of 12 (7.4%) potentially viable clinically relevant chromosomal abnormalities, which would have remained overlooked if only conventional NIPT had been performed
Luo et al. (2017)	BOLD-MRI	Propose to use BOLD-MRI with maternal hyperoxia to quantitatively assess mismatch in placental function in seven monozygotic twin pairs naturally matched for genetic growth potential	Placental function	Twin birth weights were recorded at delivery, with time from the fetal MRI scan to birth ranging from 0.2 to 5.5 weeks	Mean placental TTP negatively correlated with fetal liver and brain volumes at the time of MRI as well as with birth weights. Mean placental TTP positively correlated with placental pathology
Abaci Turk et al. (2017)	BOLD-MRI	To present a method for spatiotemporal alignment of <i>in utero</i> MRI time series acquired during maternal hyperoxia for enabling improved quantitative tracking of BOLD signal changes that characterize oxygen transport through the placenta to fetal organs	Placental functioning	Four singletons and six twin pregnancies between 26 and 34 weeks of gestational age	Unexpected temporal signal fluctuations during the first normoxia period were significantly decreased ($P < 0.01$), volume overlap and distance between region boundaries measures were significantly improved ($P < 0.01$)
Ambroise Grandjean et al. (2018)	Smartplanes® software	To evaluate the feasibility and reproducibility of AI software	Fetal health	Thirty fetuses (evaluated at 17–30 weeks' gestation), two 3D volumes of fetal head, one conventional 2D image of the transthalamic plane	High intra- and interobserver reproducibility, ICC values >0.98
Caballero-Ruiz et al. (2017)	Expectation-maximization clustering algorithm, C4.5 DT	To automatically identify the transthalamic plane from 3D ultrasound volumes and to measure the biparietal diameter and head circumference in fetus	CDSS	90 patients with GDM at Parc Tauli University Hospital	Time devoted by clinicians to patients' evaluation was reduced by 27% and face-to-face visits per patient were reduced by 89%
Moreira et al. (2018)	ANN, fuzzy logic, ANFIS	To improve access to specialized healthcare assistance, to prevent patients from unnecessary displacements, to reduce the evaluation time per patient and to avoid gestational diabetes adverse outcomes	Hypertensive disorder	205 parturient women diagnosed with a hypertensive disorder during pregnancy	AUC = 0.829

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Moreira et al. (2018)	Evolutionary radial basis function network	To identify possible cases of gestational diabetes that can lead to multiple risks for both the pregnant women and the fetus	Maternal diagnosis	National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) dataset ($n = 394$)	F-measure of 0.786
Peleg et al. (2017)	Computer-interpretable guidelines	To establish a ubiquitous, user-friendly, patient-centered mobile decision support system for patients and for their care providers, based on the continuous application of clinical guidelines and on semantically integrated electronic health records	CDSS	19 GDM patients used the GDM MobiGuide system in Spain	Achieved high compliance to self-measurement recommendations, while enhancing the satisfaction of patients and care providers
Rigla et al. (2018)	Computer-interpretable guidelines	To test the feasibility and acceptance of a mobile decision support system for GDM, developed in the seventh framework program MobiGuide Project	CDSS	20 patients diagnosed with GDM	High degree of patient satisfaction. Research demonstrates feasibility of this telemedicine application
Yang et al. (2018)	cfDNA, NIPT, SVM, RBF	To develop an accurate SVM-based approach and showed its potential in trisomy prediction on chromosomes 13/18/21	Chromosomal abnormalities	5518 NIPT data, contains 47 triploid cases	T21: predicted 100%, T18: 99.61%, T13: 99.87%
Neocleous et al. (2018)	Feed-forward ANN, cfDNA	To estimate the risk of fetal trisomy 21 (T21) and OCA at 11–13 weeks' gestation using computational intelligence classification methods	Chromosomal abnormalities	Training set = dataset consisting of 72 054 euploid pregnancies, 295 cases of T21 and 305 cases of OCA Validation = 36 058 euploid, 137 T21 and 133 OCA cases	T21 SE = 97.1%, specificity = 99.5%
Zhang et al. (2018)	Genome-wide association, multi-kernel SVM	To address the gap in our knowledge of gene regulatory elements in the placenta	Placental functioning	Trained a placental enhancer classifier using the 748 known placental enhancers as positives and a random subset of 2000 robust non-placental enhancers as the negatives	Created a map of 33 010 predicted placental enhancers, including 4562 high-confidence enhancer predictions
Looney et al. (2018)	Fully CNN (OxNNNet)	To fully automate the segmentation of an organ from 3D ultrasound (3D US) volumes, using the placenta as the target organ	Placental functioning	2393 first trimester 3D-US volumes	AUC = 0.65
Koivu et al. (2018)	SVM, DL, LR	Evaluate ML algorithms to improve performance of first trimester screening of Down syndrome	Chromosomal abnormalities	3 datasets from three individual clinical studies from routine first trimester screening programs	DNN AUC = 0.96, detection rate = 78% and 1% FPR
Moreira et al. (2018a)	ANN, MLP, PSO	The use of a biologically inspired technique, known as PSO, is proposed for reducing the computational cost of the ANN-based method referred to as the MLP, without reducing its precision rate	High-risk pregnancy	100 parturient women diagnosed with a hypertensive disorder during pregnancy	Proposed method outperformed, on average, other approaches by 26.4% in terms of precision and 14.9% in terms of the true-positive ratio (TPR), and showed a reduction of 35.4% in the FPR

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Moreira <i>et al.</i> (2018b)	Semantic interoperability, ontologies, ICD-10	To develop a knowledge-based decision support system that uses ontologies for integrating data related to hypertensive disorders in pregnancy	Pregnancy-induced hypertension	133 participants diagnosed with a pregnancy (May–September 2017)	Recent state-of-the-art approaches based on openEHR data representation are not sufficient for representing all pregnancy-related data Generalized boosted models performed best overall risk of small-for-gestational-age 1.65 fold, using the correct fetal weight standard
Naimi <i>et al.</i> (2018)	Linear and quantile regression, RF, Bayesian additive regression trees, generalized boosted models	Assess whether ML algorithms can accurately predict estimated fetal weight over the course of gestation using readily available <i>ex utero</i> information and assess the relationship between smoking and fetal/birthweight across gestation	Fetal weight	18 517 pregnancies (31 948 ultrasound visits) from the Magee-Womens Obstetric Maternal and Infant data and 240 pregnancies in a separate dataset of high-risk pregnancies, Obstetrical Determinants of Neonatal Survival (ODNSR) Simulated data	Achieves an average spatial prediction error of 7 mm on simulated data and produces qualitatively improved reconstructions for heavily moving fetuses with gestational ages of approximately 20 weeks
Hou <i>et al.</i> (2018)	CNN	Present a learning-based image registration method capable of predicting 3D rigid transformations of arbitrarily oriented 2D image slices, with respect to a learned canonical atlas co-ordinate system	Fetal health		Both ML models found an association between maternal exposure to ambient PM ₁₀ and CHDs in Beijing Discovery cohort: correctly classified patients who required CSF diversion with 82% accuracy, 80% sensitivity and 84% specificity Replication cohort: 91% accuracy, 75% sensitivity and 95% specificity Maternal serum metabolomics is a promising tool for the accurate and sensitive screening of such congenital defects Tested on 16 users, accuracy = 87.5%
Ren <i>et al.</i> (2018)	RF, gradient boosting	Investigate the non-linear effects of PM ₁₀ exposure during the critical time window, weeks 3–8 in pregnancy, on risk of congenital heart defects (CHD)	Air pollution and fetal health	39 053 live-born infants in Beijing, from 2009 to 2012	
Pisapia <i>et al.</i> (2018)	Image-based ML	To determine whether extraction of multiple imaging features from fetal MRI and integration using ML techniques can predict which patients require postnatal cerebrospinal fluid (CSF) diversion after birth	Fetal health	253 patients with fetal ventriculomegaly (2008–2014)	
Troisi <i>et al.</i> (2018)	Ensemble learning	To perform a characterization of maternal serum in order to build a metabolomic fingerprint resulting from congenital anomalies of the central nervous system	Fetal health	Metabolomic profiles were obtained from serum of 168 mothers (98 controls and 70 cases)	
Akbulut <i>et al.</i> (2018)	Averaged perceptron, boosted DT, Bayes point machine, DF, decision jungle, locally deep SVM, LR, neural network, SVM	To help clinicians and families to better predict fetal congenital anomalies besides the traditional pregnancy tests using ML techniques and e-Health applications	Fetal health	Obtained through maternal questionnaire and detailed evaluations from 3 clinicians	
Miao and Miao (2018)	DNN, ensemble learning	Developed model is used to distinguish and classify the presence or absence of multiclass morphologic patterns for outcome predictions of complications during pregnancy	Fetal health	CTG dataset with 2126 clinical instances, from 1980, and 1995–1998 (UCI Machine Learning Repository dataset)	Accuracy = 88.02%, recall = 94.30%, precision = 85.01%, F-score = 0.8508

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Vasconcelos et al. (2018)	Deep learning, ResNet101 architecture, SVM	Propose a deep learning method for classification of different surgical actions performed during laser ablation therapy during twin-to-twin transfusion syndrome interventions	Twin-to-twin transfusion syndrome (TTTS)	Videos obtained from five different TTTS ablation procedures	Automatic detection of surgical action in fetoscopic video is feasible
De Ramón Fernández et al. (2019)	ANN, MLP, DL, SVM, NB	A CDSS has been developed to help clinicians to choose the initial treatment to be followed by the patient	Ectopic pregnancy	406 cases of tubal ectopic pregnancies collected at the Department of Obstetrics and Gynecology of the University Hospital 'Virgen de la Arrixaca' in the Murcia Region (Spain) (November 2010–September 2015)	SVM and MLP show best performance SVM accuracy = 96.1%, sensitivity = 96%, specificity = 98%
Al-Shaer et al. (2019)	Principal component analysis, k-means clustering	To discover signaling pathways and miRNAs that inform mechanisms underlying fetal alcohol spectrum disorder (FASD)	Fetal alcohol spectrum disorder	RNA-seq dataset derived from neural fold-stage chicken embryos that were exposed to pharmacologically relevant alcohol concentration or isotonic saline, followed by a 4.5-hour recovery period. Cranial headfolds were isolated 6 hours following initial exposure	Identified 6857 differentially expressed exons representing 1251 geneIDs Identified exons encoding 23 microRNAs having significantly differential expression profiles in response to alcohol
Burton et al. (2019)	Random forest, neural network, Extreme Gradient Boosting	Explored the potential for a ML solution to reduce the burden of culturing the large number of culture-negative samples without reducing detection of culture-positive samples	Maternal diagnosis	212 554 urine reports	Workload reduction of 41% and a sensitivity of 95%
Corona-Figueroa (2019)	ANFIS, rule-based expert system	Develop a prototype meant to be an auxiliary for the detection of fetal arrhythmias by analyzing the FHR and its variability	Fetal heart health	Set of 5 artificial signals Set of real ECG signals from a group of 9 pregnant women from 8 electrodes at a sample rate of 500 Hz	Synthetic signal tests, DR = 88.88%
Gao et al. (2019)	Observational Medical Outcomes Partnership (OMOP) common data model, ICD, CPT, regularized logistic regression models	To identify severe maternal morbidity (SMM) during delivery hospitalizations, we introduce a ML framework that incorporates data in EHRs	Maternal morbidity	45 858 deliveries for obstetric inpatients, from Vanderbilt University Medical Center EHR system (2005–2008)	AUC = 0.94, 85% patients diagnosed with fluid and/or electrolytes disorder experienced SMM during their delivery hospitalization, 82% patients diagnosed with systemic inflammatory response syndrome developed SMM
Gomes Filho et al. (2019)	Bayesian networks, rule-based	To present a hybrid model capable of performing the early diagnosis of GDM based on the disease code, analyzing the medical history of pregnant patients	Maternal diagnosis	The database of a health insurance company which covers eleven Brazilian states (January 2004–December 2009)	Prototype demonstrates feasibility
Mohseni Salehi et al. (2019)	CNN	To increase the capture range and accelerate the performance of state-of-the-art inter-subject and subject-to-template 3D rigid registration	Fetal brain	93 reconstructed T2-weighted MRI scans of fetuses, as well as T1- and T2-weighted MRI scans of 40 newborns	Proposed methods provided very fast (real-time) registration with a wide capture range on the space of all plausible 3D rotations and provided good initialization

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Moriera et al. (2019)	Averaged one-dependence estimators (AODE) algorithm, Bayes' theorem, ANN, DT	The use of AODE is proposed for real-time pregnancy data analysis from Internet of Things (IoT) devices and gateways	Pregnancy complications	205 parturient women diagnosed with a hypertensive disorder during pregnancy	AODE algorithm has inferior accuracy performance (versus DT and ANNs), but superior computational time performance
Signorini et al. (2020)	ApEn, Lempel Ziv complexity, phase-rectified signal average, robust linear regression, RF, SVM, classification trees, elastic NET, LASSO, NB	Tested the performance of 15 ML techniques in discriminating healthy versus IUGR fetuses	Intrauterine growth restriction (IUGR)	CTG traces and prenatal fetal condition from a database of 60 healthy and 60 IUGR fetuses	RFs achieved best performance Mean accuracy for all 10 tests = 0.911
Lu et al. (2020)	Ensemble model, random forest, XGBoost, LightGBM, genetic algorithm	To predict fetal weight at varying gestational age in the absence of ultrasound examination within a certain accuracy	Fetal weight	4212 intrapartum recordings	Ensemble model: accuracy = 64.3%, mean relative error = 7.0%
Preterm Grignolio (1982)	Multiple regression	Predict probability of being alive at 30 days of life for premature newborns weighing less than 2000 g	Neonate mortality	Weight of 92 preterm neonates	80% accuracy on prediction
Woolery (1994)	Multiple regression analysis, rule-based (ID3 algorithm)	Develop clinical knowledge base for preterm birth risk assessment	Preterm risk	2739/~16 000 cases from a database of high-risk pregnancy women referred for home uterine monitoring services	Overall, experts determined rules did not provide enough information, missing important data
Wu et al. (2004)	Adaptive speckle suppression using variable length sticks algorithm, region-based segmentation, curve-fitting function	Develop an algorithm that is able to identify the anatomic landmarks of the cervix on a transvaginal ultrasound image and determine the standard cervical length	Cervical length	101 cervical ultrasound images were selected from a series of 37 examinations performed on 17 patients over an 8-month period	Pairwise comparison, the mean cervical length for the computer method matches with the mean manual cervical length
Alamedine et al. (2013)	Jeffrey divergence (JD) distance, a sequential forward selection (SFS) and a binary PSO (BPSO), quadratic discriminant analysis (QDA), k-NN, LDA	To reduce the number of features from EHG signals by selecting only the relevant ones which are useful for solving the classification problem	Fetal health	Signals recorded on 48 women: 32 during pregnancy (33–39 weeks of gestation) and 16 during labor (39–42 weeks of gestation)	Feature selected by BPSO with QDA corresponds to the highest percentage of correct classification (88.72%)
Fergus et al. (2013)	Synthetic minority oversampling technique (SMOTE), density-based classifiers, linear- and polynomial-based classifiers, non-linear-based classifiers	Presents a supervised ML approach that classifies term and preterm records	Fetal health	Open-source dataset containing 300 records (38 preterm and 262 term).	96% sensitivity, 90% specificity and a 95% area under the curve value with 8% global error using the polynomial classifier

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Góralaska et al. (2013)	CHAID classification tree	To prove that precise information about the proper course of action that is included in the expanded Apgar score might have a prognostic value for the early death incidence in preterm newborns	Fetal health	276 children (born from <32 weeks of gestational age pregnancies) at the 2nd Department of Obstetrics and Gynecology, who were hospitalized between the years 2009 and 2011 at the Neonatology and Intensive Care Unit, Medical University of Warsaw	Concentration of oxygen used during resuscitation at birth (a component of the expanded Apgar score) has significantly more prognostic value
Cerqueira et al. (2014)	ANN, SVM, cost-sensitive algorithm	Describes NICESim, an open-source simulator that uses ML techniques to aid health professionals to better understand the treatment and prognosis of premature newborns	Fetal health	Data collected in the Neonatal Intensive Care Unit of São Sebastião Hospital, at Viçosa, Minas Gerais State, Brazil, (2008–2010), 293 instances	Accuracy = 86.7%
Sadi-Ahmed et al. (2017)	Huang-Hilbert transform, intrinsic mode function, SVM	To achieve a fully automated method for predicting preterm deliveries only basing on EHG signal analysis	Preterm birth prediction	30 pregnant women (1997–2005), 15 of which delivered prematurely, EHG records (27–32 weeks of gestation), 1 record per patient	Two-channel combination achieved the best mean results with accuracy = 95.70%, sensitivity = 98.40%, specificity = 93.00%, AUC = 0.95.
Bahado-Singh et al. (2019)	DL, RF, SVM, LDA, GLM, prediction analysis for microarrays	To evaluate the application of AI to amniotic fluid metabolomics and proteomics, alone and in combination with sonographic, clinical and demographic factors, in the prediction of perinatal outcome in asymptomatic pregnant women with short cervical length	Preterm birth prediction	32 cases, Early preterm (<34 weeks), control (≥ 34 weeks)	DL showed best performance, AUC = 0.890 for delivery <34 weeks' gestation, 0.890 for delivery <28 days post-amniocentesis and 0.792 for NICU admission
Lee and Ahn (2019)	ANN, DT, NB, RF, SVM, multinomial LR	Analyze preterm birth and its determinants based on general hospital data and compare ANN to other ML methods	Preterm birth preventive measures	596 obstetric patients during 27 March 2014–21 August 2018 at Anam Hospital in Seoul, Korea	Accuracy of ANN comparable to LR and RF Determinants = BMI, hypertension, DM, prior cone biopsy, prior placenta previa, parity, cervical length, age, prior preterm birth, and myomas and adenomyosis Expected error of 6 days and a 30% improvement over prediction based on weight only
Torres Torres et al. (2019)	SVM, GNN	Present a novel method for postnatal gestational age estimation that eliminated the subjectivity issues present in the Ballard Score	Postnatal gestational age prediction	The GestAtional Dataset (2015–2017), 130 babies' images of foot, ear and face	
Gao et al. (2019)	DL, recurrent neural networks (RNN)	Investigated the extent to which DL models that consider temporal relations documented in EHRs can predict extreme preterm birth (EPB)	Preterm birth prediction	EHR data from 5689 deliveries at Vanderbilt University Medical Center (9 August 2005–31 July 2017)	RNN ensemble model: AUC = 0.82, sensitivity = 0.96
Shi et al. (2019)	Genetic algorithm	To measure the anisotropic mechanical properties of human cervical tissue using indentation and video extensometry	Cervix tissue properties	Material properties of human cervical slices taken at different anatomical locations from 7 women of different obstetric backgrounds	The anatomic anterior and posterior quadrants of cervical tissue are more anisotropic than the left and right quadrants. There is no significant difference in material properties between samples of different parities

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Perinatal, birth and delivery Hernández <i>et al.</i> (1986)	Syntactic pattern recognition combined with Hon's Apgar prediction method	To overcome communication of automatic pattern recognition to obstetric professionals with a topogram	FHR and prognosis with Apgar prediction	Uterine pressure via intrauterine catheters FHR via subcutaneous electrodes Tested on 69 patients	95 ± 1% success at predicting Apgar rating at birth
Anderson and Barclay (1988)	Date conversion and averages	Allow clinician to input up to 20 clinical data points and determine average estimated day of delivery	Estimated date of delivery	Not validated	N/A
Sokol and Chik (1998)	Text processing utilities from UNIX, coded into PROLOG compatible files	Expert system for perinatal knowledge	Clinical knowledge system	Prototype system	N/A
Alonso-Betanzos <i>et al.</i> (1991)	Mixed frame- and rule-based approach for clinical knowledge representation ANN	Help clinical professionals interpret: biophysical profile, contraction stress test and non-stress test. It also interprets supervision of labor and newborn diagnosis Evaluate an intelligent system to assist clinicians in their assessment of fetal condition and management of labor by expert and system agreement	Fetal health	87 cases randomly selected from Ob-Gyn Dept. of the Hospital General de Galicia, Spain	Retrospective and prospective validation—similar performance to expert systems of the time
Keith <i>et al.</i> (1994)	ANN	Evaluate an intelligent system to assist clinicians in their assessment of fetal condition and management of labor by expert and system agreement	Fetal health	Intrapartum CTG and patient information 100 training set, 30 test set	System achieved good agreement with 3 experts
Beksaçet <i>et al.</i> (1995)	Supervised ANN (back propagation learning algorithm)	Development of an AI diagnostic system for the interpretation of umbilical artery blood flow velocity waveform measurements	Fetal health	Training set: 952 umbilical artery blood flow velocity waveform images of 174 normal pregnancies Testing set: 138 images derived from 20 normal pregnancies, 68 images of 16 high risk pregnancies	Specificity = 98.6% Sensitivity = 51.5%
Beksaçet <i>et al.</i> (1996)	Generalized learning vector quantization (GLVQ) algorithm	To develop an automated intelligent diagnostic system for the interpretation of umbilical artery velocity waveforms	Fetal health	Training set: 278 and 380 waveform images of 94 normal and 157 high risk pregnancies Testing set: 193 normal pregnancy and 61 high-risk pregnancy waveform images	Specificity = 80.3% Sensitivity = 54.1%
Devoue <i>et al.</i> (1996)	Neural network (three-layered, feed-forward with a back propagation)	To create a neural network that predicts the length of the first stage of term labor	Labor management	159 for training and 41 for testing	36 of 41 first-stage durations were predicted within 2 hours of actual length

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Ungureanu et al. (2007)	Maternal beat subtraction with linear modeling, noise cancelling methods,	Propose a new method to subtract mECG after detecting and removing abdominal signal segments with high-amplitude variations due to uterine contractions	Fetal health	Signals were recorded at a sampling rate of 400 Hz, using unipolar measurement and 12 electrodes were placed on the abdomen. Fetal ECG was also acquired using fetal scalp electrodes at a sampling rate of 1024 Hz	Comparable performance to methods of the time
La Rosa et al. (2008)	Schwarz information criterion (SIC), non-supervised k-means cluster algorithm	Propose a single-channel two-stage time-segment discriminator of uterine magnetomyogram (MMG) contractions during pregnancy	Fetal health	MMG records obtained from 10 patients admitted to the hospital for contractions (31–40 weeks of gestation)	Detection rate and false alarm rate are favorable
Terrien et al. (2008)	Time-frequency representation, gradient vector flow	Describe a new ridge extraction method based on the image processing technique of active contours or snakes	Fetal health	Testing: two pregnant monkeys by the implantation of a telemetric transducer	Application on real EHG: able to obtain 63 contractions per placental location for a total of 126 contractions and identified a total of 52 polyphasic contractions
Yilmaz and Kilikçier (2013)	Least-squares SVM, PSO, binary DT	To classify CTG data to determine the fetal state as normal, suspect or pathologic	Fetal health	CTG data set = 2126 data points, with 21 features, taken from the UCI Machine Learning Repository	Accuracy = 91.62%
Liu et al. (2013)	Genetic algorithm	To identify novel serological protein markers to diagnose preeclampsia (PE) with a multi 'omics' -based discovery approach	Maternal diagnosis	n = 111 PE and n = 152 control placenta samples	Developed two optimal biomarker panels for early and late onset PE assessment
Ge et al. (2014)	Spatial-color pixel classification, maximizing cluster validity index, k-NN	Describes a system for automated counting and distinguishing fetal and maternal RBCs on clinical KB slides	Fetal-maternal hemorrhage (FMH)	120 raw color images of KB slides	Capable of completing the counting of over 60 000 cells (versus ~2000 by technologists) within 5 minutes (versus ~15 minutes by technologists)
Sahin and Subasi (2015)	SVM, ANN, radial basis functions, DTs, k-NN, RF	To evaluate the classification performances of eight different ML methods on the antepartum cardiocogram (CTG) data	Fetal health	UCI CTG dataset with 1831 instances, each containing 21 features (8 continuous, 13 discrete)	RF can be accepted as a good classifier of normal and pathological classes of the CTG data
Czabanski et al. (2016)	Fuzzy inference-based learning, Lagrangian SVM	Present a method of retrospective fetal state assessment using the results of the fuzzy analysis of delivery outcome attributes	Fetal health	CTG recordings from 1-hour fetal monitoring sessions via an external pulsed Doppler ultrasound transducer	accuracy = 99.18% Fuzzy reasoning improved the efficiency of the LSVM classification when compared with the standard learning
Chen and Hao (2017)	Extreme learning machine, single-hidden layer feedforward neural network, Hilbert-Huang transform	Present a method for feature extraction and classification of EHG between pregnancy and labor group	Labor classification	122 EHG recordings (112 pregnancy recordings and 10 labor recordings) made on 45 pregnant women	Accuracy = 88.00%, sensitivity = 91.30%, specificity = 85.19%

(Continued)

Table 1. Continued

Study	Methods	Objective	Health issue	Dataset	Result
Alberola-Rubio <i>et al.</i> (2017)	SVM, majority voting	To help distinguish patients that will achieve labor spontaneously from those that will need late-term induction of labor	Route of delivery	EHG recordings from 62 singleton term pregnancies, 38 spontaneously entered labor and 24 induced	Aggregation of majority voting of SVM models using contextualized EHG parameters: AUC = 0.93
Fergus <i>et al.</i> (2017)	DL, RF, recursive feature eliminator algorithm, SMOTE, Fishers LDA	Present a proof-of-concept using ML and FHR signals as an ambulatory decision support to antenatal care	Route of delivery	506 normal vaginal deliveries and 46 cases of cesarean section, CTG recordings for singleton pregnancies, gestational age <36 weeks (CTU-UHB database)	DL classifier, sensitivity = 94%, specificity = 91%, AUC = 99%, F-score = 100%, and MSE = 1%
Beksac <i>et al.</i> (2018)	ANN	To develop a supportive computerized prediction system for the route of delivery	Route of delivery	2127, 3548 and 1723 deliveries for the years 1976, 1986 and 1996	Specificity = 97.5%, Sensitivity = 60.9%, FPR = 2.5%, FNR = 39.1%, PPV = 81.8%, NPV = 93.1%
Fergus <i>et al.</i> (2018)	Synthetic minority oversampling, Fishers LDA (FDLA), RF, SVM	To determine whether ML could be used to optimize the decisions made by obstetricians and midwives, in the presence of FHR traces, linked to caesarean section and normal vaginal deliveries	Route of delivery	552 records from the CTU-UHB database	Ensemble classifier: sensitivity = 87%, specificity = 80%, AUC = 96%, MSE = 8%
Postnatal and postpartum Green <i>et al.</i> (2019b)	mHealth, NLP, Tess AI, DT	To test the Healthy Moms perinatal depression intervention using a single-case experimental design with pregnant women and new mothers recruited from public hospitals outside of Nairobi, Kenya	Perinatal depression	Initial prepilot round of formative testing with 10 women from a private maternity hospital in May and June	Prototype followed up with more research
Wang <i>et al.</i> (2019)	L2-regularized LR, SVM, DT, NB, XGBoost, RF	To demonstrate that ML models can be used to predict PPD and to carefully evaluate the risk factors identified from EHR data	Postpartum depression	9980 episodes of pregnancy (EHRs from Weill Cornell Medicine and New York-Presbyterian Hospital between 2015 and 2017)	SVM: AUC = 0.79, Significant predictors = race, obesity, anxiety, depression, different types of pain, antidepressants and anti-inflammatory drugs during pregnancy
Any stage Prieto <i>et al.</i> (2014)	SVMs, NB, DT, k-NN	Exploit the wealth of user-generated data, available through the microblogging service Twitter, to estimate and track the incidence of health conditions in society	Social media	More than 10 million tweets, Spanish and Portuguese	Correlation-based feature selection proved to be the most effective
Yom-Tov <i>et al.</i> (2015)	ICD-10, linear SVM, the Self-Controlled Case Series method	Report that Web search engine queries coupled with information on Wikipedia access patterns can be used to infer health events associated with an individual user and automatically generate Web-based risk markers for some of the common medical conditions worldwide (including pregnancy)	Internet usage	English language queries submitted to the Bing search engine by users in the United States for the 6-month period starting May 2012, zip code of user and small proportion of users provided year of birth	Most common medical condition identified was pregnancy
Zhang and Demner-Fushman (2017)	SVM, NLP	To develop automated classification methods for eligibility criteria in ClinicalTrials.gov to facilitate patient-trial matching for specific populations such as persons living with HIV or pregnant women	Clinical trials	891 interventional cancer trials and the eligibility criteria for 1570 of the most recent interventional trials from ClinicalTrials.gov for HIV-positive and pregnancy eligibility	Achieved macro-averaged F2 scores of 0.91 and 0.85 for HIV and pregnancy, relatively
Klein <i>et al.</i> (2018)	NLP, rule-based, boot-strapping	To assess whether rare health-related events (birth defects) are reported on social media, to design an NLP approach for collecting data from social media and to utilize the collected data to discover a cohort of women whose pregnancies with birth defect outcomes could be observed for epidemiological analysis	Social media	Mined 432 million tweets posted by 112 647 users who were automatically detected via their public announcements of pregnancies on Twitter	Discovery of 195 users that met the inclusion criteria Recall = 0.95

Table 2. ML methods and their applications to the pregnancy domain

ML techniques		Application to pregnancy	
Method	Description of method	Pregnancy area most used	Informatics domain
Supervised			
SVM	Constructs a hyperplane or set of hyperplanes; finds the decision boundary for classification.	<ul style="list-style-type: none"> • ART 	<ul style="list-style-type: none"> • Prediction and assessment
ANN	Composed of nodes or states that are connected to a second set of nodes via functions; contains an input layer, at least one hidden layer and an output layer	<ul style="list-style-type: none"> • Preterm birth • Prenatal • Perinatal, labor and delivery • Postpartum • ART 	<ul style="list-style-type: none"> • Chromosomal anomalies • CDSS • Electronic monitoring • NLP and search analytics • Prediction and assessment
MLP	A class of feedforward ANN.	<ul style="list-style-type: none"> • Prenatal • Preterm birth • Perinatal, labor and delivery 	<ul style="list-style-type: none"> • Chromosomal anomalies • CDSS • mHealth • Electronic monitoring • Prediction and assessment • CDSS • Prediction and assessment
Radial basis function (RBF) network	A type of ANN that performs classification by measuring the input's similarity to examples from the training set	<ul style="list-style-type: none"> • ART • Prenatal • Perinatal, labor and delivery 	<ul style="list-style-type: none"> • Chromosomal anomalies • Prediction and assessment
RF	A classification method that contains multiple decision trees	<ul style="list-style-type: none"> • Preconception 	<ul style="list-style-type: none"> • Chromosomal anomalies • Prediction and assessment
NB	A family of algorithms based on Bayes' theorem with an assumption of independence between predictors	<ul style="list-style-type: none"> • Prenatal • Preterm birth • Postpartum • Preconception 	<ul style="list-style-type: none"> • CDSS • Radiology, Imaging • NLP and search analytics • Prediction and assessment
k-NN	An algorithm that classifies new cases based on a similarity measure	<ul style="list-style-type: none"> • ART • Prenatal • Postpartum • Art 	<ul style="list-style-type: none"> • CDSS • mHealth • NLP and search analytics • Chromosomal anomalies

(Continued)

Table 2. Continued

Method	ML techniques		Application to pregnancy	
	Description of method	Additional references for more details on method	Pregnancy area most used	Informatics domain
Discriminant analysis	A method that finds a set of prediction equations based on independent variables that are used to classify individuals into groups	[162]	<ul style="list-style-type: none"> • Perinatal, labor and delivery • Preterm birth 	<ul style="list-style-type: none"> • Electronic monitoring • Prediction and assessment
Genetic algorithm (GA)	An algorithm that imitates the process of natural selection	[163]	<ul style="list-style-type: none"> • Perinatal, labor and delivery • Art 	<ul style="list-style-type: none"> • Genetics and biomarkers • Prediction and assessment
Semi-supervised DL	Constructs of ANN using a large number hidden layers	[164]	<ul style="list-style-type: none"> • Prenatal screening • Preterm birth • Perinatal, labor and delivery 	<ul style="list-style-type: none"> • Genetics and biomarkers
Unsupervised Principle component analysis	Constructs of ANN using a large number hidden layers	[164]	<ul style="list-style-type: none"> • Art 	<ul style="list-style-type: none"> • Prediction and assessment
Independent component analysis	A dimensionality reduction method for large datasets, often used before the ML algorithm. It attempts to find uncorrelated components	[165]	<ul style="list-style-type: none"> • Prenatal • Preterm • Perinatal, labor and delivery 	<ul style="list-style-type: none"> • Radiology, Imaging
k-means clustering	A technique for decomposing a complex dataset into independent components	[166]	<ul style="list-style-type: none"> • Prenatal • Perinatal care, labor and delivery • Prenatal 	<ul style="list-style-type: none"> • Electronic monitoring • Genetics and biomarkers
	A clustering method that partitions data into a specified number of K clusters	[167]	<ul style="list-style-type: none"> • Perinatal care, labor and delivery • Prenatal 	<ul style="list-style-type: none"> • Electronic monitoring • Genetics and biomarkers
			<ul style="list-style-type: none"> • Perinatal, labor and delivery 	<ul style="list-style-type: none"> • Electronic monitoring

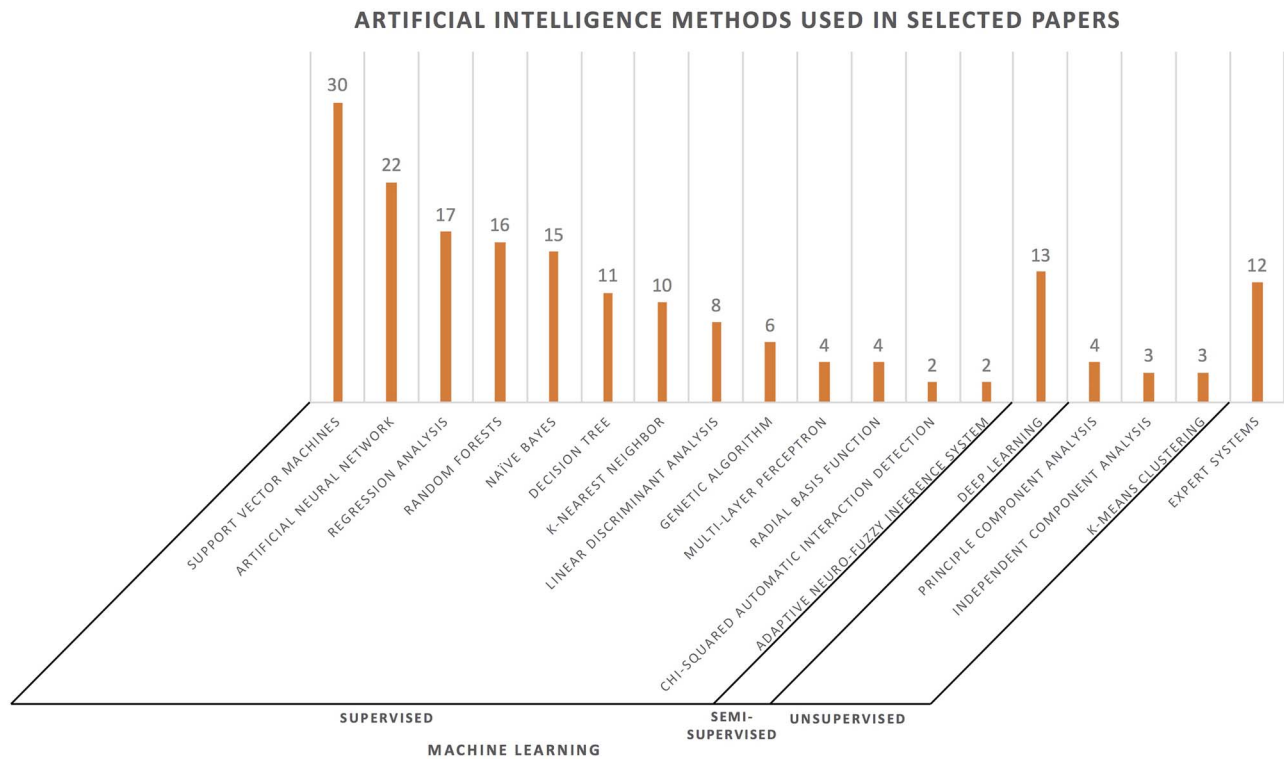


Figure 3. AI methods used in studies included in review. ML methods are ordered by frequency and further categorized by supervised and unsupervised learning methods. Totals here do not equal the number of included papers, as studies often demonstrate several methods.

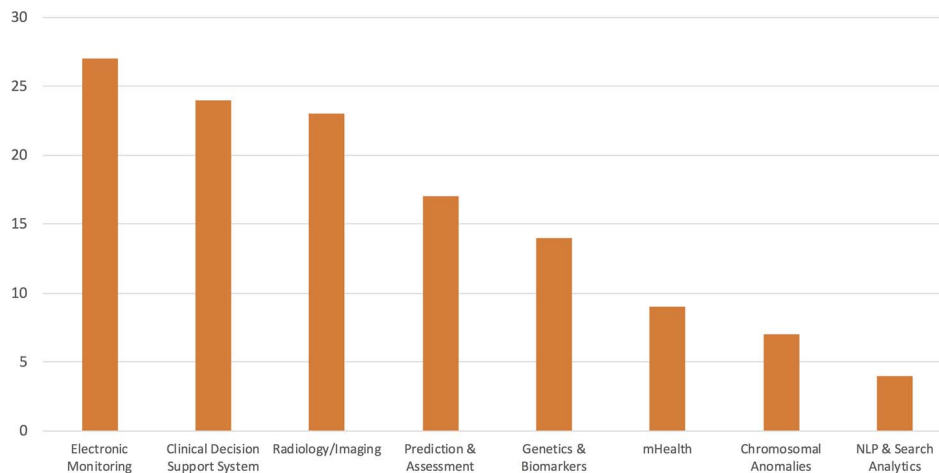


Figure 4. Informatics areas utilized by included studies. Because a study may utilize multiple domains (e.g. NLP + Radiology), there may be overlap across domains for a given study. For example, mHealth applications overlap multiple domains.

pregnancy care, especially for pregnancy with comorbidities [39, 40], will be needed to provide adequate care while reducing exposure during prenatal care.

NLP and search analytics

A wealth of data is accessible through social media, including health behaviors and health outcomes. NLP techniques were applied to user-generated Twitter data to estimate and track the incidence of pregnancy [41] and collect data on birth defect outcomes [42]. Support vector machine (SVM) was used to infer health events associated with individual

users and found pregnancy to be the most commonly identified medical condition [43]. NLP techniques and SVM were applied to automatically classify eligibility criteria in [ClinicalTrials.gov](https://clinicaltrials.gov) to facilitate patient-trial matching [44]. A great tool for public health and pharmacovigilance, NLP can analyze unstructured health data, improve EHR usability, and facilitate interoperability.

Assisted reproductive technology

AI assessment of embryo images or videos has great potential to improve ART outcomes. Applications guide identification of embryos from the culture medium during early human

in vitro development [45], raw time-lapse videos/images of embryos [46–48] and in vitro fertilization (IVF) EHR data [49, 50]. Evaluation of morphokinetic time-lapse microscopy data depends on the experience and knowledge of embryologists—this work is highly subjective with a lack of standardization. Applications may reveal details of embryo morphology imperceptible to the human eye or predict successful pregnancy by integrating other relevant health data points. DL can determine robust quantitative imaging biomarkers for embryo selection, improve therapy outcomes and reduce clinical burden.

Assisted Reproductive Technology CDSS

A convolutional neural network (CNN) was implemented to select highest quality embryos using a large collection of human embryo time-lapse images from a high-volume fertility center in the United States [48]. This 22-layer deep model resulted in the trained algorithm called STORK. STORK performed well on additional datasets of embryo images from two other IVF centers. Different grading systems (i.e. unstandardized) across clinics affect performance; STORK demonstrated lower performance for one clinic dataset and therefore lower generalizability. For embryo selection, STORK outperformed individual embryologists in assessing embryo image quality. However, STORK could not predict positive and negative live births using embryo morphology alone.

Medical imaging in pregnancy

Fetal development

Applications of AI have targeted fetal development: predicting fetal health status [36], improving fetal brain imaging [51–53] and other fetal anatomy [54, 55]. Semi-automated learning-based framework approach reported to improve gestational age prediction, with accuracy of ± 6.1 days, using DF on structural brain image and clinical data [56]. AI has been applied to improve knowledge and treatment of ectopic pregnancy, using ML classifiers [57] and using gene stability algorithms [58]. Studies aim to improve imaging of fetal organ development with virtual organ computer-aided analysis (VOCAL) [55], texture analysis [52] and CNN [53, 59]. Three studies observe EHR data; higher performing classification ML approaches used in these studies include DF [36, 56] and SVM [57].

Placental functioning

Placental health is deeply connected with maternal–fetal health and plays an essential role in supporting fetal development [60]. Placental function and characteristics have been observed with VOCAL with 3D power Doppler, blood oxygen level-dependent magnetic resonance imaging (BOLD-MRI) [61, 62], ultrasound images [63] and CNN [64]. Research focuses on improving placental pathology and sufficiency knowledge [62], placental volume imaging methods [61, 65] and placental maturity identification [63]. BOLD-MRI has gained attention as a promising non-invasive technique to monitor placental function *in vivo*; however, the use in practice is limited.

Diagnostic imaging CDSS

The prenatal diagnosis rate of major congenital heart defects remains low most likely due to the downfalls of manual navigation during sonographic screening, as it is operator-dependent, challenging and time-consuming. The Fetal Intelligent Navigation Echocardiography (FINE) method provides visualization

of standard fetal electrocardiogram (fECG) views from volume datasets obtained with spatiotemporal image correlation, aiming to consistently display diagnostic planes, regardless of fetal position or initial orientation [54]. The Virtual Intelligent Sonographer Assistant tool was developed to visualize specific structures in the complex anatomy of the fetal heart. When applied to four abnormal cases, the FINE method demonstrated evidence of abnormal fetal cardiac anatomy in multiple fECG views.

Genetic screening and chromosomal abnormalities

Definitive prenatal diagnosis of genetic disease is performed via invasive procedures, specifically amniocentesis or chorionic villus sampling. Risks include miscarriage, fetal morbidity, parental anxiety and Rh sensitization [66]. Non-invasive prenatal tests (NIPTs) hold no such risks and usually tests only for trisomy 21 (T21). Trisomies 18 (T18) and 13 (T13) are the second and third most common trisomies after T21. Expansion of NIPT could include screening for T18 and T13 [67, 68] and other chromosomal abnormalities (OCA) [69, 70]. A number of ML techniques were applied to improve chromosomal screening, including artificial neural network (ANN) [69, 70]; SVM [70, 71]; *k*-nearest neighbor [70, 71] and deep NN, RF, NB, DT and logistic regression (LR) [71]. Studies comparing methods found NN methods to perform best [70, 71]. AI methods were applied to optimize detection of relatively small mutations at a low sequencing cost [72]. In order to build a ‘metabolic fingerprint’ resulting from congenital anomalies of the central nervous system, ensemble learning was applied to characterize maternal serum [73].

During pregnancy, the pregnant person and the fetal-placental unit produce cell-free DNA (cfDNA). Using next generation sequencing, a SVM model demonstrated high accuracy for cfDNA testing [74]. Genome-wide cfDNA analysis detected OCA with high sensitivity, in comparison with standard cfDNA screening [75]. A proposed two-stage routine procedure, with combined tests, invasive tests and cfDNA tests, achieved high detection rate for T21 cases and proved to be minimally invasive and of relatively low cost [76].

Chromosomal screening CDSS

Nuchal translucency (NT) measurement, a well-established ultrasonographic marker for fetal aneuploidy screening, can vary greatly and as a result impair screening performance. The performance-adjusted risk (PAR) method allows for these differences in measurements and improves performance [77]. PAR analyzes individual provider and laboratory marker distribution parameters, compares with national expectations and then assigns a handicap. NT and serum markers are considered in the method, using commercially available software and knowledge from meta-analysis of published literature. The PAR method informs providers of their handicap and their performance in relationship to others.

Electronic monitoring

The FHR signal is more complicated to identify than the adult signal, and therefore, there has been an effort to improve methods to detect, sample and quantify the FHR signal accurately. Computer analysis of the FHR is categorized into three stages: (1) raw signal processing, (2) pattern analysis and (3) expert/intelligent systems [23]. Early computer programs only identified deceleration patterns and outcomes of normal, warning or ominous patterns. Studies apply AI methods to ECG

[78–84], magnetocardiography (MCG) [85, 86] and cardiogram (CTG) [87–90]. In attempt to improve fetal signal extraction from ECG and MCG recordings, several supervised learning methodologies were proposed: adaptive neuro-fuzzy inference (ANFIS) [79, 84] and time–frequency analysis [81]. Unsupervised learning techniques were applied: independent component analysis [78, 83, 85, 86], principle component analysis [46, 91] and Kalman filtering [82]. Moreover, studies applied a variety of ML methods to classify CTG signals and determine fetal state, with best performances demonstrated using SVM [87, 89], DL [90] and RF [88].

Fetal monitoring and diagnostic CDSS

A portable prototype expert system detected modeled fetal arrhythmias with 88% accuracy and could be applied to scenarios where there is an insufficient number of experts and a large number of patients [84]. The prototype uses ANFIS to extract fECG from thoracic and abdominal signals, extracts the FHR and then detects fetal arrhythmia through an expert system based on production rules.

Pregnancy complications

Maternal diagnosis during pregnancy

Accurate prediction methods and diagnosis during prenatal care allow health professionals and prospective parents detect problems with the pregnancy as early as possible. Some studies investigate patient comorbidities including gestational diabetes mellitus [18, 27, 92–97], gestational hypertension disorders [98–100] and bacteriuria [101]. Studies aim to predict and classify disease in early pregnancy, improve screenings and provide clinical decision support for disease management. Methods proposed include rule-based [27, 97], Bayesian networks [18, 94, 95], ANN [92, 98, 101], evolutionary radial basis function network [93], genetic algorithm [99], expert systems [94, 96] and DT [96].

Preterm birth

Concerning preterm birth, studies focus on cervix-related risk [102, 103], classifying true preterm labor [104, 105], determining neonate mortality and prognosis [106, 107], predicting preterm birth risk [108–111], professional learning [112], estimating post-natal gestational age [113] and improving knowledge of gene regulatory elements in the placenta [114]. In 1982, Grignolio [106] presented a method to predict neonate mortality of premature newborns, using multiple regression. In 1990, Andersen et al. [115] found that shorter cervical length (CL) was associated with a high risk of preterm delivery. Woolery [108, 116] developed a clinical knowledge base for preterm birth risk assessment and developed an expert system for preterm birth risk assessment of pregnant women.

For pregnancy and labor contraction classification, studies found that AI methods improve diagnosis from EHG records, including polynomial classifiers [104], feature selection using binary particle swarm optimization with quadratic discriminant analysis [105] and a ridge extraction method [117]. Cervix properties were observed using AI methods to determine material properties [103] and CL [115]. Perinatal outcome was predicted with a DL model interpreting amniotic fluid metabolomics and proteomics in asymptomatic pregnant women with short CL [109]. SVM was applied to predict genome-wide placental enhancers, in order to further the understanding

of placental dysfunction and implications of preterm birth and preeclampsia [114].

Two studies turn to EHR data to determine preterm birth risk, using ANN, LR, RF [110] and recurrent neural networks [111]. Both studies found associations with preterm birth risk and hypertensive disorders, as well as CL.

CDSS for pregnancy complications

Ectopic pregnancy can be potentially life-threatening, and therefore, early diagnosis and treatment of the condition are needed. A three-stage classifier (3SC) developed finds and limits diagnostic errors and then assists clinicians in choosing the initial treatment of ectopic pregnancy [57]. The model was evaluated across four algorithms: SVM, NB, DL and auto multilayer perceptron (MLP). 3SC achieved the best performance, in comparison with single-stage classifiers.

At the event of fetal–maternal hemorrhage, the standard clinical method to quantify fetal and maternal red blood cells (RBCs) is the Kleihauer–Betke test, which is performed by a certified technologist. The automated system can count over 60 000 RBCs within 5 minutes, versus a technologist counting ~2000 in about 15 minutes [118].

The Apgar score [119, 120] evaluates the condition of newborns in their first minutes of neonatal life. However, the score is not intended as a prognostic value; it has limitations and is often highly subjective, depending largely on the clinical experience of the evaluator. The prognostic value of the Expanded Apgar Score Form [121] was studied, observing the early death incidence in preterm newborns [107]. The chi-square automatic interaction detection classification tree generated decision rules based upon clinical data. The concentration of the oxygen applied during resuscitation was found to be an important criterion to neonatal outcome.

An open-source simulator, called NICeSim, was developed to aid health professionals to better understand the treatment and prognosis of premature newborns admitted to the neonatal intensive care unit [112]. The system allows for flexibility and provides alterations of the chosen variables. The attributes used to calculate death risk include Apgar scale, respiratory distress syndrome, gestational age and birth weight. Sufficient predictive power was demonstrated; ANN achieved better accuracy and specificity, whereas SVM performed better sensitivity. Important attributes are most likely absent as the ML model uses four attributes for a complex problem.

Perinatal care, birth and delivery

Fetal weight estimation

Clinicians and prospective parents want to be informed about the well-being of the unborn infant. AI applications observe infant growth and health status from imaging and maternal EHR data history. ML methods were applied to predict babies' birth weight, using LR [122, 123], MLP [123] and fuzzy logic support vector regression [124], RF [122], Bayesian models [122, 125] and generalized boosted model [122]. The studies predict weight using EHR data [123, 124, 126], amniotic fluid [125], ultrasound images [122] and CTG traces [127].

Contraction classification

When extracting the FHR from an ECG recording, uterine contractions during labor can introduce significant noise. In order to improve perinatal knowledge and outcomes, AI has been applied to estimating date of delivery [115], route of delivery

[128–131], fetal health prognosis [132] and length of labor [133]. Methods were developed to provide perinatal knowledge [134], improve fetal monitoring and assessment [87–89, 117, 135–139] and classify true labor contractions [140]. A single paper employed unsupervised ML, using k-means clustering [139]. A 2015 systematic review found no strong evidence that the use of CTG with an expert system has an effect on the incidence of cesarean delivery nor a reduction in the incidence of forceps-assisted vaginal birth [24].

Route of delivery

Solely based upon FHR traces, models were developed to improve classifying route of delivery [129, 131] and classifying pregnancy and labor contractions [105]. Performance varied between ML techniques and number of features included in models; using 13 features, the DL classifier demonstrated the best performance, while selecting 8 features, the RF classifier performed best [129]. An ensemble classifier of three ML techniques, FLDA, RF and SVM classifiers, allows the strength of each model to be used to classify between cesarean section and vaginal delivery types [131].

Models take advantage of available EHR data in conjunction with labor monitoring data. Another study found that an aggregation double-layer SVM model using contextualized EHG parameters performed best to distinguish patients that will achieve spontaneous labor before the end of their full term from those that will require late-term induction of labor based upon EHG signals and obstetrical parameters [128]. Obstetrical parameters include maternal age, BMI, gestations, parity, Bishop score and days of gestation at recording moment.

CDSS for birth and delivery

The Adana System applied ANN to classify between cesarean section and vaginal delivery, including input variables of maternal characteristics and labor information [130]. Predicting route of delivery can inform care, allow appropriate allocation of resources and improve pregnancy outcomes.

Postpartum

During the postpartum period, many mothers try to cope with physical, social and psychological changes [141, 142]. Postpartum depression (PPD) affects an estimated 13–19% of women who have recently given birth [143]. Wang *et al.* [144] applied several ML models to predict PPD using EHRs and found several associated risk factors: race, obesity, anxiety, depression, antidepressants and anti-inflammatory drugs during pregnancy and different types of pain [144].

Unintended consequences and AI ethics

The potential of AI for pregnancy care exists but not without its barriers and potential pitfalls. Challenges include dataset shift, discriminatory bias, generalizability, poor clinical applicability, accidental fitting of cofounders and unintended negative consequences on health outcomes [145].

For instance, no papers were retrieved concerning transgender and gender non-conforming pregnancy. We also performed another literature search and found that there were no transgender-related AI studies on pregnancy. Many transgender men and gender non-conforming individuals (assigned female at birth) retain the capacity to become pregnant, use contraception, desire to become pregnant and give birth, and there

are some case studies on this process [146, 147]. Little research focuses on the reproductive needs of this population [146, 148–150] and stigma can lead individuals to avoid seeking medical care or disclosing medically relevant health information [151]. Transgender individuals are likely to have unique needs surrounding pregnancy health and psychosocial health (e.g. gender dysphoria, lactation following chest surgery or binding, hormone replacement therapy) [151]. It would not be ethical to employ off-the-shelf AI methods designed for gender-conforming populations to those that were transgender or gender non-conforming populations.

It is unrealistic to assume that AI applications in health are ethically neutral. AI technology has the capacity to violate the basic rights of individuals, such as learning patterns from past clinical care that might increase discrimination (e.g. structural racism), violate privacy and inadvertently create inequalities in care. Unintentional discrimination (e.g. vulnerable populations, sexism and racism) need to be proactively accounted for in algorithms [152, 153], especially given that algorithms utilize data collected in the past that may be unintentionally biased in various ways. For patients and healthcare providers, individuals need a good understanding of the algorithm's decision-making in order to have agency over the derived clinical decisions. In congruence with current best practices, AI needs to provide room for healthcare providers to influence the decision-making process. The patterns underlying AI decisions should also be made transparent to researchers for them to understand the potential biases underlying these algorithms.

Limitations and future work

We used 'pregnancy' as one of the search terms for identifying relevant papers. Therefore, preconception and postpartum studies were only included if they also discussed a pregnancy element. This may have lowered the overall number of studies focused on preconception and postpartum stages of pregnancy included in our literature review. Preconception care and postpartum care encompass healthcare beyond the scope of this particular review, and therefore, this study does not represent AI and ML applications in the entirety of these fields. Studies included in our review do not explicitly report the gender of their patients; therefore, it is unclear how many transgender patients may have been included in the research studies that we cite.

Conclusion

We found that AI and ML methodologies are used to help inform pregnancy outcomes. Of which 16 took advantage of EHR data to derive sound and reliable information about maternal and fetal health—illustrating the need to increase research using EHR data. Applications of mHealth were used in nine studies in the following areas: preconception [26, 33], during gestation [27, 36, 93, 96, 97, 154] and postpartum [34]. Unsupervised methods were limited to automated EHR and contraction monitoring [78, 82, 83, 85, 86, 139], imaging [46], genetics [91] and trisomy screening [69] applications. Research aimed to detect anomalies and diagnoses earlier, inform practice and support patients, improve fetal and maternal health knowledge, deep phenotype obstetric complications and automate aspects of current practice.

Past reviews of AI applications in pregnancy research and healthcare have described great potential but little realization in the clinical setting. While the papers in our review did not describe adoption explicitly in the clinical setting, the variety of CDSS ($n=24$) and mHealth applications ($n=9$) in our review

illustrate the effort to translate AI into clinical care. Low adoption of AI in clinical care in the pregnancy domain may likely be due to liability questions in case of medical error that remain unresolved—especially for so-called black-box algorithms. More research on maternal healthcare is needed, with only 31% of prenatal stage papers focused on maternal needs. In addition, future research should proactively account for unintentional discrimination in algorithms. Potential pitfalls need to be accounted for in development of AI algorithms, in order to prevent adverse events and negative clinical impacts. Greater use of unsupervised methods may be appropriate with only nine studies using this approach but could be limited by the interpretability of their results. DL has begun to gain traction in recent years [10]—unsupervised ML and DL naturally apply well to imaging analysis and large datasets in biomedicine. The amalgamation of supervised and unsupervised learning (i.e. semi-supervised ML) could be valuable, as labeling data for supervised ML methods is time-consuming and expensive. AI and ML methods can be successfully employed to optimize pregnancy outcomes; with proper algorithmic improvement, refinement and ethics results, these methods could be incorporated into clinical care.

Key Points

- Artificial intelligence and machine learning methods can be successfully employed to optimize pregnancy outcomes.
- Supervised methods are the most common approach, and deep learning methods are gaining traction.
- Future research should focus on less-studied areas: maternal healthcare needs, postpartum care, pregnancy care for transgender individuals and in algorithmic improvement, refinement and translation of results back into clinical care.

Supplementary data

Supplementary data are available online at *Briefings in Bioinformatics*.

Data Availability Statement

This is a review paper and the primary data is the analysis of the papers themselves. This analysis and the main result of our work is included in [Table 1](#) for all to access. We have also included a Prisma checklist as a supplementary document. Any additional data requests could be sent to the corresponding author: bolandm@upenn.edu

Funding

This work was supported by the Perelman School of Medicine at the University of Pennsylvania with generous funds to support this project.

Authors' contribution

L.M.D. and M.R.B. conceived study design; wrote paper and reviewed, edited and approved final manuscript.

References

1. Shahid N, Rappon T, Berta W. Applications of artificial neural networks in health care organizational decision-making: a scoping review. *PLoS One* 2019;**14**:e0212356.
2. Fatima M, Pasha M. Survey of machine learning algorithms for disease diagnostic. *J Intell Learn Syst Appl* 2017;**09**:1–16.
3. Robinson PN. Deep phenotyping for precision medicine. *Hum Mutat* 2012;**33**:777–80.
4. Weng C, Shah NH, Hripcsak G. Deep phenotyping: embracing complexity and temporality—towards scalability, portability, and interoperability. *J Biomed Inform* 2020;**105**:103433.
5. Sen C. Preterm labor and preterm birth. *J Perinat Med* 2017;**45**:911–3.
6. Correa PJ, Palmeiro Y, Soto MJ, et al. Etiopathogenesis, prediction, and prevention of preeclampsia. *Hypertens Pregnancy* 2016;**35**:280–94.
7. Larsen EC, Christiansen OB, Kolte AM, et al. New insights into mechanisms behind miscarriage. *BMC Med* 2013;**11**:154.
8. Deo RC. Machine learning in medicine. *Circulation* 2015;**132**:1920–30.
9. Handelman GS, Kok HK, Chandra RV, et al. eDoctor: machine learning and the future of medicine. *J Intern Med* 2018;**284**:603–19.
10. Miotto R, Wang F, Wang S, et al. Deep learning for healthcare: review, opportunities and challenges. *Brief Bioinform* 2018;**19**:1236–46.
11. IMIA & Georg Thieme Verlag KG Stuttgart. Contents IMIA yearbook of medical informatics 2019. *Yearb Med Inform* 2019;**28**:I–VIII.
12. Raef B, Ferdousi R. A review of machine learning approaches in assisted reproductive technologies. *Acta Inform Med* 2019;**27**:205–11.
13. Balayla J, Shrem G. Use of artificial intelligence (AI) in the interpretation of intrapartum fetal heart rate (FHR) tracings: a systematic review and meta-analysis. *Arch Gynecol Obstet* 2019;**300**:7–14.
14. Iftikhar PM, Kuijpers MV, Khayyat A, et al. Artificial intelligence: a new paradigm in obstetrics and Gynecology research and clinical practice. *Cureus* 2020;**12**:e7124.
15. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;**6**:e1000100.
16. Williamson PO, Minter CIJ. Exploring PubMed as a reliable resource for scholarly communications services. *J Med Libr Assoc* 2019;**107**:16–29.
17. Adlassnig K-P. *Medical Informatics in a United and Healthy Europe: Proceedings of MIE 2009, the XXII International Congress of the European Federation for Medical Informatics*, 2009, 238. IOS Press, Amsterdam, Netherlands.
18. Hernando ME, Gómez EJ, Corcoy R, et al. Evaluation of DIABNET, a decision support system for therapy planning in gestational diabetes. *Comput Methods Prog Biomed* 2000;**62**:235–48.
19. Catanzarite VA, Jelovsek FR. Computer applications in obstetrics. *Am J Obstet Gynecol* 1987;**156**:1049–53.
20. Singh AV, Rouchka EC, Rempala GA, et al. Integrative database management for mouse development: systems and concepts. *Birth Defects Res Part C Embryo Today Rev* 2007;**81**:1–19.

21. Greene KR. Intelligent fetal heart rate computer systems in intrapartum surveillance. *Curr Opin Obstet Gynecol* 1996;**8**:123–7.
22. Devoe LD. Future perspectives in intrapartum fetal surveillance. *Best Pract Res Clin Obstet Gynaecol* 2016;**30**:98–106.
23. McCartney PR. Computer analysis of the fetal heart rate. *J Obstet Gynecol Neonatal Nurs* 2000;**29**:527–36.
24. Lutomski JE, Meaney S, Greene RA, et al. Expert systems for fetal assessment in labour. *Cochrane Database Syst Rev* 2015;**4**:1–2.
25. Haddad SM, Souza RT, Cecatti JG. Mobile technology in health (mHealth) and antenatal care—searching for apps and available solutions: a systematic review. *Int J Med Inform* 2019;**127**:1–8.
26. Goodale BM, Shilaih M, Falco L, et al. Wearable sensors reveal menses-driven changes in physiology and enable prediction of the fertile window: observational study. *J Med Internet Res* 2019;**21**:e13404.
27. Rigla M, Martínez-Sarriegui I, García-Sáez G, et al. Gestational diabetes management using smart mobile telemedicine. *J Diabetes Sci Technol* 2018;**12**:260–4.
28. Benski AC, Stancanelli G, Scaringella S, et al. Usability and feasibility of a mobile health system to provide comprehensive antenatal care in low-income countries: PANDA mHealth pilot study in Madagascar. *J Telemed Telecare* 2017;**23**:536–43.
29. World Health Organization. *OpenSRP|Open Smart Register Platform*. 2020.
30. Stroux L, Martinez B, Coyote Ixen E, et al. An mHealth monitoring system for traditional birth attendant-led antenatal risk assessment in rural Guatemala. *J Med Eng Technol* 2016;**40**:356–71.
31. Bakibinga P, Kamande E, Omuya M, et al. The role of a decision-support smartphone application in enhancing community health volunteers' effectiveness to improve maternal and newborn outcomes in Nairobi, Kenya: quasi-experimental research protocol. *BMJ Open* 2017;**7**:e014896.
32. Marko KI, Krapf JM, Meltzer AC, et al. Testing the feasibility of remote patient monitoring in prenatal care using a mobile app and connected devices: a prospective observational trial. *JMIR Res Protoc* 2016;**5**:e200.
33. Green EP, Whitcomb A, Kahumbura C, et al. 'What is the best method of family planning for me?': a text mining analysis of messages between users and agents of a digital health service in Kenya. *Gates Open Res* 2019;**3**:1475.
34. Green EP, Pearson N, Rajasekharan S, et al. Expanding access to depression treatment in Kenya through automated psychological support: protocol for a single-case experimental design pilot study. *JMIR Res Protoc* 2019;**8**:e11800.
35. Cunningham SD, Lewis JB, Thomas JL, et al. Expect with me: development and evaluation design for an innovative model of group prenatal care to improve perinatal outcomes. *BMC Pregnancy Childbirth* 2017;**17**:147.
36. Akbulut A, Ertugrul E, Topcu V. Fetal health status prediction based on maternal clinical history using machine learning techniques. *Comput Methods Prog Biomed* 2018;**163**:87–100.
37. Lim J, Cloete G, Dunsmuir DT, et al. Usability and feasibility of PIERS on the move: an mHealth app for pre-eclampsia triage. *JMIR Mhealth Uhealth* 2015;**3**:e37.
38. Balakrishnan R, Gopichandran V, Chaturvedi S, et al. Continuum of care services for maternal and child health using mobile technology - a health system strengthening strategy in low and middle income countries. *BMC Med Inform Decis Mak* 2016;**16**:1–8.
39. Murphy HR. Managing diabetes in pregnancy before, during, and after COVID-19. *Diabetes Technol Ther* 2020;**22**:454–61.
40. Aziz A, Zork N, Aubey JJ, et al. Telehealth for high-risk pregnancies in the setting of the COVID-19 pandemic. *Am J Perinatol* 2020;**37**:800–8.
41. Prieto VM, Matos S, Álvarez M, et al. Twitter: a good place to detect health conditions. *PLoS One* 2014;**9**:e86191.
42. Klein AZ, Sarker A, Cai H, et al. Social media mining for birth defects research: a rule-based, bootstrapping approach to collecting data for rare health-related events on Twitter. *J Biomed Inform* 2018;**87**:68–78.
43. Yom-Tov E, Borsa D, Hayward AC, et al. Automatic identification of web-based risk markers for health events. *J Med Internet Res* 2015;**17**:e29.
44. Zhang K, Demner-Fushman D. Automated classification of eligibility criteria in clinical trials to facilitate patient-trial matching for specific patient populations. *J Am Med Assoc* 2017;**24**:781–7.
45. Nyalwidhe J, Burch T, Bocca S, et al. The search for biomarkers of human embryo developmental potential in IVF: a comprehensive proteomic approach. *Mol Hum Reprod* 2013;**19**:250–63.
46. Milewski R, Kuczyńska A, Stankiewicz B, et al. How much information about embryo implantation potential is included in morphokinetic data? A prediction model based on artificial neural networks and principal component analysis. *Adv Med Sci* 2017;**62**:202–6.
47. Tran D, Cooke S, Illingworth PJ, et al. Deep learning as a predictive tool for fetal heart pregnancy following time-lapse incubation and blastocyst transfer. *Hum Reprod* 2019;**34**:1011–8.
48. 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;**21**:1–8.
49. Morales DA, Bengoetxea E, Larrañaga P, et al. Bayesian classification for the selection of in vitro human embryos using morphological and clinical data. *Comput Methods Prog Biomed* 2008;**90**:104–16.
50. Uyar A, Bener A, Ciray HN. Predictive modeling of implantation outcome in an in vitro fertilization setting: an application of machine learning methods. *Med Decis Mak* 2015;**35**:714–25.
51. Ambroise Grandjean G, Hossu G, Bertholdt C, et al. Artificial intelligence assistance for fetal head biometry: assessment of automated measurement software. *Diagn Interv Imaging* 2018;**99**:709–16.
52. Gentillon H, Stefańczyk L, Strzelecki M, et al. Parameter set for computer-assisted texture analysis of fetal brain. *BMC Res Notes* 2016;**9**:496.
53. Mohseni Salehi SS, Khan S, Erdogmus D, et al. Real-time deep pose estimation with geodesic loss for image-to-template rigid registration. *IEEE Trans Med Imaging* 2019;**38**:470–81.
54. Yeo L, Romero R. Fetal intelligent navigation echocardiography (FINE): a novel method for rapid, simple, and automatic examination of the fetal heart. *Ultrasound Obstet Gynecol* 2013;**42**:268–84.
55. Re C, Bertucci E, Weissmann-Brenner A, et al. Fetal thymus volume estimation by virtual organ computer-aided analysis in normal pregnancies. *J Ultrasound Med* 2015;**34**:847–52.

56. Namburete AIL, Stebbing RV, Kemp B, et al. Learning-based prediction of gestational age from ultrasound images of the fetal brain. *Med. Image Anal* 2015;21:72–86.
57. De Ramón FA, Ruiz Fernández D, Prieto Sánchez MT. A decision support system for predicting the treatment of ectopic pregnancies. *Int J Med Inform* 2019;129:198–204.
58. Gebeh AK, Marczylo EL, Amoako AA, et al. Variation in stability of endogenous reference genes in fallopian tubes and endometrium from healthy and ectopic pregnant women. *Int J Mol Sci* 2012;13:2810–26.
59. Hou B, Khanal B, Alansary A, et al. 3-D reconstruction in canonical co-ordinate space from arbitrarily oriented 2-D images. *IEEE Trans Med Imaging* 2018;37:1737–50.
60. Burton GJ, Fowden AL. The placenta: a multifaceted, transient organ. *Philos Trans R Soc B Biol Sci* 2015;370:20140066.
61. Turk EA, Luo J, Gagoski B, et al. Spatiotemporal alignment of in utero BOLD-MRI series. *J Magn Reson Imaging* 2017;46:403–12.
62. Luo Y, Li Z, Guo H, et al. Predicting congenital heart defects: a comparison of three data mining methods. *PLoS One* 2017;12:e0177811.
63. Lei B, Yao Y, Chen S, et al. Discriminative learning for automatic staging of placental maturity via multi-layer fisher vector. *Sci Rep* 2015;5:12818.
64. Looney P, Stevenson GN, Nicolaides KH, et al. Fully automated, real-time 3D ultrasound segmentation to estimate first trimester placental volume using deep learning. *JCI Insight* 2018;3:e120178.
65. Sanderson J, Wu L, Mahajan A, et al. Selection of the sub-noise gain level for acquisition of VOCAL data sets: a reliability study. *Ultrasound Med Biol* 2014;40:562–7.
66. Centers for Disease Control and Prevention. Chorionic villus sampling and amniocentesis: recommendations for prenatal counseling. *MMWR Recomm Rep* 1995;44:1–12.
67. Kagan KO, Wright D, Valencia C, et al. Screening for trisomies 21, 18 and 13 by maternal age, fetal nuchal translucency, fetal heart rate, free β -hCG and pregnancy-associated plasma protein-a. *Hum Reprod* 2008;23:1968–75.
68. Marttala J, Peuhkurinen S, Ranta JK, et al. Screening and outcome of chromosomal abnormalities other than trisomy 21 in northern Finland. *Acta Obstet Gynecol Scand* 2011;90:885–9.
69. Neocleous AC, Nicolaides KH, Schizas CN. Intelligent noninvasive diagnosis of aneuploidy: raw values and highly imbalanced dataset. *IEEE J Biomed Health Informatics* 2017;21:1271–9.
70. Neocleous AC, Nicolaides KH, Schizas CN. First trimester noninvasive prenatal diagnosis: a computational intelligence approach. *IEEE J Biomed Health Informatics* 2016;20:1427–38.
71. Koivu A, Korpimäki T, Kivelä P, et al. Evaluation of machine learning algorithms for improved risk assessment for Down's syndrome. *Comput Biol Med* 2018;98:1–7.
72. Chu T, Yeniterzi S, Rajkovic A, et al. High resolution non-invasive detection of a fetal microdeletion using the GCREM algorithm. *Prenat Diagn* 2014;34:469–77.
73. Troisi J, Landolfi A, Sarno L, et al. A metabolomics-based approach for non-invasive screening of fetal central nervous system anomalies. *Metabolomics* 2018;14:1–10.
74. Yang J, Ding X, Zhu W. Improving the calling of non-invasive prenatal testing on 13-/18-/21-trisomy by support vector machine discrimination. *PLoS One* 2018;13:1–19.
75. Fiorentino F, Bono S, Pizzuti F, et al. The clinical utility of genome-wide non-invasive prenatal screening. *Prenat Diagn* 2017;37:593–601.
76. Neocleous AC, Syngelaki A, Nicolaides KH, et al. Two-stage approach for risk estimation of fetal trisomy 21 and other aneuploidies using computational intelligence systems. *Ultrasound Obstet Gynecol* 2018;51:503–8.
77. Evans MI, Cuckle HS. Performance adjusted risks: a method to improve the quality of algorithm performance while allowing all to play. *Prenat Diagn* 2011;31:797–801.
78. Meinecke F, Ziehe A, Kawanabe M, et al. A resampling approach to estimate the stability of one-dimensional or multidimensional independent components. *IEEE Trans Biomed Eng* 2002;49:1514–25.
79. Assaleh K. Extraction of Fetal electrocardiogram using adaptive neuro-fuzzy inference systems. *IEEE Trans Biomed Eng* 2007;54:59–68.
80. Kumar, Sharma SK, Prasad S. CAD for detection of fetal electrocardiogram by using wavelets and neuro-fuzzy systems. *Int J Appl Eng Res* 2016;11:2321–6.
81. Karvounis EC, Tsipouras MG, Fotiadis DI, et al. An automated methodology for Fetal heart rate extraction from the abdominal electrocardiogram. *IEEE Trans Inf Technol Biomed* 2007;11:628–38.
82. Haghpanahi M, Borkholder DA. Fetal QRS extraction from abdominal recordings via model-based signal processing and intelligent signal merging. *Physiol Meas* 2014;35:1591–605.
83. Varanini M, Tartarisco G, Billeci L, et al. An efficient unsupervised fetal QRS complex detection from abdominal maternal ECG. *Physiol Meas* 2014;35:1607–19.
84. Corona-Figueroa A. *A portable prototype for diagnosing fetal arrhythmia*. Unlocked: Informatics Med, 2019.
85. Mantini D, Alleva G, Comani S. A method for the automatic reconstruction of fetal cardiac signals from magnetocardiographic recordings. *Phys Med Biol* 2005;50:4763–81.
86. Comani S, Srinivasan V, Alleva G, et al. Entropy-based automated classification of independent components separated from fMCG. *Phys Med Biol* 2007;52:N87–97.
87. Yilmaz E, Kiliççier Ç. Determination of fetal state from cardiocogram using LS-SVM with particle swarm optimization and binary decision tree. *Comput Math Methods Med* 2013;2013:1–8.
88. Sahin H, Subasi A. Classification of the cardiocogram data for anticipation of fetal risks using machine learning techniques. *Appl Soft Comput J* 2015;33:231–8.
89. Czabanski R, Jezewski M, Horoba K, et al. Fuzzy analysis of delivery outcome attributes for improving the automated Fetal state assessment. *Appl Artif Intell* 2016;30:556–71.
90. Miao JH, Miao KH. Cardiocographic diagnosis of fetal health based on multiclass morphologic pattern predictions using deep learning classification. *Int J Adv Comput Sci Appl* 2018;9:1–11.
91. Al-Shaer AE, Flentke GR, Berres ME, et al. Exon level machine learning analyses elucidate novel candidate miRNA targets in an avian model of fetal alcohol spectrum disorder. *PLoS Comput Biol* 2019;15:1–25.
92. Polak S, Mendyk A. Artificial intelligence technology as a tool for initial GDM screening. *Expert Syst Appl* 2004;26:455–60.
93. Moreira MWL, Rodrigues JJPC, Kumar N, et al. Evolutionary radial basis function network for gestational diabetes data analytics. *J Comput Sci* 2018;27:410–7.

94. Gomes Filho E, Pinheiro PR, Pinheiro MCD, et al. Heterogeneous methodology to support the early diagnosis of gestational diabetes. *IEEE Access* 2019;7:67190–9.
95. Hernando ME, Gómez EJ, Del Pozo F, et al. DIABNET: a qualitative model-based advisory system for therapy planning in gestational diabetes. *Med Inf* 1996;21:359–74.
96. Caballero-Ruiz E, García-Sáez G, Rigla M, et al. A web-based clinical decision support system for gestational diabetes: automatic diet prescription and detection of insulin needs. *Int J Med Inform* 2017;102:35–49.
97. Peleg M, Shahar Y, Quaglini S, et al. Assessment of a personalized and distributed patient guidance system. *Int J Med Inform* 2017;101:108–30.
98. Tejera E, Joseareias M, Rodrigues A, et al. Artificial neural network for normal, hypertensive, and preeclamptic pregnancy classification using maternal heart rate variability indexes. *J Matern Neonatal Med* 2011;24:1147–51.
99. Liu LY, Yang T, Ji J, et al. Integrating multiple ‘omics’ analyses identifies serological protein biomarkers for preeclampsia. *BMC Med* 2013;11:236.
100. Moreira MWL, Rodrigues JJPC, Al-Muhtadi J, et al. Neuro-fuzzy model for HELLP syndrome prediction in mobile cloud computing environments. *Concurrency Computat Pract Exper*. 2018;e4651.
101. Burton RJ, Albur M, Eberl M, et al. Using artificial intelligence to reduce diagnostic workload without compromising detection of urinary tract infections. *BMC Med Inform Decis Mak* 2019;19:171.
102. Wu M, Fraser RF, Chen CW. A novel algorithm for computer-assisted measurement of cervical length from transvaginal ultrasound images. *IEEE Trans Inf Technol Biomed* 2004;8:333–42.
103. Shi L, Yao W, Gan Y, et al. Anisotropic material characterization of human cervix tissue based on indentation and inverse finite element analysis. *J Biomech Eng* 2019;141:0910171–3.
104. Fergus P, Cheung P, Hussain A, et al. Prediction of preterm deliveries from EHG signals using machine learning. *PLoS One* 2013;8:e77154.
105. Alamedine D, Khalil M, Marque C. Comparison of different EHG feature selection methods for the detection of preterm labor. *Comput Math Methods Med* 2013;2013:585–93.
106. Grignolio G. Medical diagnoses by artificial intelligence process. *Clin Exp Obstet Gynecol* 1982;9:233–4.
107. Górska A, Puskarz-Gąsowska J, Borszewska-Kornacka MK. Practical application and prognostic value of the expanded Apgar score. *Arch Perinat Med* 2013;19:88–93.
108. Woolery LK, Grzymala-Busse J. Machine learning for an expert system to predict preterm birth risk. *J Am Med Inform Assoc* 1994;1:439–46.
109. Bahado-Singh RO, Sonek J, McKenna D, et al. Artificial intelligence and amniotic fluid multiomics: prediction of perinatal outcome in asymptomatic women with short cervix. *Ultrasound Obstet Gynecol* 2019;54:110–8.
110. Lee KS, Ahn KH. Artificial neural network analysis of spontaneous preterm labor and birth and its major determinants. *J Korean Med Sci* 2019;34:e128.
111. Gao C, Osmundson S, Velez Edwards DR, et al. *Deep learning predicts extreme preterm birth from electronic health records*. Inform: J Biomed, 2019.
112. Cerqueira FR, Ferreira TG, de Paiva OA, et al. NiCeSim: an open-source simulator based on machine learning techniques to support medical research on prenatal and perinatal care decision making. *Artif Intell Med* 2014;62:193–201.
113. Torres Torres M, Valstar M, Henry C, et al. Postnatal gestational age estimation of newborns using small sample deep learning. *Image Vis Comput* 2019;83–84:87–99.
114. Zhang J, Simonti CN, Capra JA. Genome-wide maps of distal gene regulatory enhancers active in the human placenta. *PLoS One* 2018;13:e0209611.
115. Andersen HF, Nugent CE, Wanty SD, et al. Prediction of risk for preterm delivery by ultrasonographic measurement of cervical length. *Am J Obstet Gynecol* 1990;163:859–67.
116. Fitzgerald Woolery LK. Clinical knowledge base development for preterm-birth risk assessment. *Appl Nurs Res* 1994;7:141–4.
117. Terrien J, Marque C, Germain G. Ridge extraction from the time–frequency representation (TFR) of signals based on an image processing approach: application to the analysis of uterine electromyogram AR TFR. *IEEE Trans Biomed Eng* 2008;55:1496–503.
118. Ge J, Gong Z, Chen J, et al. A system for counting fetal and maternal red blood cells. *IEEE Trans Biomed Eng* 2014;61:2823–9.
119. Apgar V. A proposal for a new method of evaluation of the newborn. *Class Pap Crit Care* 1952;32:97.
120. Apgar V, Holaday DA, James LS, et al. Evaluation of the newborn infant-second report. *JAMA* 1958;168:1985–8.
121. Stark AR, Adamkin DH, Batton DG, et al. The Apgar score. *Pediatrics* 2006;117:1444–7.
122. Naimi AI, Platt RW, Larkin JC. Machine learning for fetal growth prediction. *Epidemiology* 2018;29:290–8.
123. Etikan I, Çağlar MK. Prediction methods for babies’ birth weight using linear and nonlinear regression analysis. *Technol Health Care* 2005;13:131–5.
124. Yu J, Wang Y, Chen P. Fetal weight estimation using the evolutionary fuzzy support vector regression for low-birth-weight fetuses. *IEEE Trans Inf Technol Biomed* 2009;13:57–66.
125. Boisvert MR, Koski KG, Burns DH, et al. Early prediction of macrosomia based on an analysis of second trimester amniotic fluid by capillary electrophoresis. *Biomark Med* 2012;6:655–62.
126. Lu Y, Fu X, Chen F, et al. Prediction of fetal weight at varying gestational age in the absence of ultrasound examination using ensemble learning. *Artif Intell Med* 2020;102:101748.
127. Signorini MG, Pini N, Malovini A, et al. Integrating machine learning techniques and physiology based heart rate features for antepartum fetal monitoring. *Comput Methods Prog Biomed* 2020;185:105015.
128. Alberola-Rubio J, Garcia-Casado J, Prats-Boluda G, et al. Prediction of labor onset type: spontaneous vs induced; role of electrohysterography? *Comput Methods Prog Biomed* 2017;144:127–33.
129. Fergus P, Hussain A, Al-Jumeily D, et al. Classification of caesarean section and normal vaginal deliveries using foetal heart rate signals and advanced machine learning algorithms. *Biomed Eng Online* 2017;16:89.
130. Beksac MS, Tanacan A, Bacak HO, et al. Computerized prediction system for the route of delivery (vaginal birth versus cesarean section). *J Perinat Med* 2018;46:881–4.
131. Fergus P, Selvaraj M, Chalmers C. Machine learning ensemble modelling to classify caesarean section and vaginal delivery types using cardiocography traces. *Comput Biol Med* 2018;93:7–16.
132. Hernández C, Arjas JE, Gómez L. A perinatal monitoring display based on the fetal topogram. *IEEE Trans Biomed Eng* 1986;BME-33:785–92.

133. Devoe LD, Samuel S, Prescott P, et al. Predicting the duration of the first stage of spontaneous labor using a neural network. *J Matern Neonatal Med* 1996;5:256–61.
134. Sokol RJ, Chik L. A prototype system for perinatal knowledge engineering using an artificial intelligence tool. *J Perinat Med* 1988;16:273–81.
135. Keith RDF, Westgate J, Hughes GW, et al. Preliminary evaluation of an intelligent system for the management of labour. *J Perinat Med* 1994;22:345–50.
136. Beksaç MS, Egemen A, Izzetoğlu K, et al. An automated intelligent diagnostic system for the interpretation of umbilical artery Doppler velocimetry. *Eur J Radiol* 1996;23:162–7.
137. Beksaç MS, Başaran F, Eskiizmirli S, et al. A computerized diagnostic system for the interpretation of umbilical artery blood flow velocity waveforms. *Eur J Obstet Gynecol Reprod Biol* 1996;64:37–42.
138. Ungureanu M, Bergmans JWM, Oei SG, et al. Fetal ECG extraction during labor using an adaptive maternal beat subtraction technique. *Biomed Tech* 2007;52:56–60.
139. La Rosa PS, Nehorai A, Eswaran H, et al. Detection of uterine MMG contractions using a multiple change point estimator and the K-means cluster algorithm. *IEEE Trans Biomed Eng* 2008;55:453–67.
140. Chen L, Hao Y. Discriminating pregnancy and labour in electrohysterogram by sample entropy and support vector machine. *J Med Imaging Health Informatics* 2017;7:584–91.
141. Aber C, Weiss M, Fawcett J. Contemporary women's adaptation to motherhood: the first 3 to 6 weeks postpartum. *Nurs Sci Q* 2013;26:344–51.
142. Aydemir S, Onan N. The relationship between maternal self-confidence and postpartum depression in primipara mothers: a follow-up study. *Community Ment Health J* 2020;56:1–8.
143. O'Hara MW, McCabe JE. Postpartum depression: current status and future directions. *Annu Rev Clin Psychol* 2013;9:379–407.
144. Wang S, Pathak J, Zhang Y. Using electronic health records and machine learning to predict postpartum depression. *Stud Health Technol Inform* 2019;264:888–92.
145. Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019;17:195.
146. Light AD, Obedin-Maliver J, Sevelius JM, et al. Transgender men who experienced pregnancy after female-to-male gender transitioning. *Obstet Gynecol* 2014;124:1120–7.
147. Light A, Wang LF, Zeymo A, et al. Family planning and contraception use in transgender men. *Contraception* 2018;98:266–9.
148. MacDonald T, Noel-Weiss J, West D, et al. Transmasculine individuals' experiences with lactation, chestfeeding, and gender identity: a qualitative study. *BMC Pregnancy Childbirth* 2016;16:106.
149. Obedin-Maliver J, Makadon HJ. Transgender men and pregnancy. *Obstet Med* 2016;9:4–8.
150. Sterling J, Garcia MM. Fertility preservation options for transgender individuals. *Transl Androl Urol* 2020;9:S215–26.
151. Hoffkling A, Obedin-Maliver J, Sevelius J. From erasure to opportunity: a qualitative study of the experiences of transgender men around pregnancy and recommendations for providers. *BMC Pregnancy Childbirth* 2017;17:332.
152. Kearns M, Roth A. The ethical algorithm: the science of socially aware algorithm design. New York, NY: Oxford University Press, 2019.
153. The Lancet Digital Health. There is no such thing as race in health-care algorithms. *Lancet Digit Health* 2019;1:e375.
154. Moreira MWL, Rodrigues JJPC, Furtado V, et al. Averaged one-dependence estimators on edge devices for smart pregnancy data analysis. *Comput Electr Eng* 2019;77:435–44.
155. Schölkopf B. SVMs - a practical consequence of learning theory. *IEEE Intell Syst Their Appl* 1998;13:18–21.
156. Renganathan V. Overview of artificial neural network models in the biomedical domain. *Bratislava Med J* 2019;120:536–40.
157. Pal SK, Mitra S. Multilayer perceptron, fuzzy sets, and classification. *IEEE Trans Neural Netw* 1992;3:683–97.
158. Dash CSK, Behera AK, Dehuri S, et al. Radial basis function neural networks: a topical state-of-the-art survey. *Open Comput Sci* 2016;6:33–63.
159. Dankowski T, Ziegler A. Calibrating random forests for probability estimation. *Stat Med* 2016;35:3949–60.
160. Puga JL, Krzywinski M, Altman N. Points of significance: Bayes' theorem. *Nat Methods* 2015;12:277–8.
161. Keller JM, Gray MR. A fuzzy k-nearest neighbor algorithm. *IEEE Trans Syst Man Cybern* 1985;SMC-15:580–5.
162. Friedman JH. Regularized discriminant analysis. *J Am Stat Assoc* 1989;84:165–75.
163. Mitchell M. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press, 2020.
164. Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface* 2018;15:20170387.
165. Bro R, Smilde AK. Principal component analysis. *Anal Methods* 2014;6:2812–31.
166. Tharwat A. Independent component analysis: an introduction. *Appl. Comput. Inform.* 2018.
167. Jain AK. Data clustering: 50 years beyond k-means. *Pattern Recogn Lett* 2010;31:651–66.