Research article

Machine learning-based evaluation of application value of the USM combined with NIPT in the diagnosis of fetal chromosomal abnormalities

Xianfeng Xu[1], **Liping Wang**[1], **Xiaohong Cheng**[2], **Weilin Ke**[2], **Shenqiu Jie**[2], **Shen Lin**[3], **Manlin Lai**[3], **Linlin Zhang**[3] **and Zhenzhou Li**[3,*]

[1] Department of Reproductive Medicine, Shenzhen Second People's Hospital, Guangdong Province, 518035, China

[2] Department of Obstetrics and Gynecology, Shenzhen Second People's Hospital, Guangdong Province, 518035, China

[3] Department of Ultrasonic Diagnosis, Shenzhen Second People's Hospital, Guangdong Province, 518035, China

* **Correspondence:** Email: lizhenzhou2004@126.com; Tel: +86-15986630196; Fax: +1-0755-83356952.

Abstract: Objective: To explore the soft ultrasound marker (USM) combined with non-invasive prenatal testing (NIPT) in diagnosing fetal chromosomal abnormalities based on machine learning and data mining techniques. **Methods**: To analyze the data of ultrasonic examination from 856 cases with high-risk single pregnancy during early and middle pregnancy stage. NIPT was applied in 642 patients. All 856 patients accepted amniocentesis and chromosome karyotype analysis to determine the efficacy of USM, Down's syndrome screening, and NIPT in detecting fetal chromosomal abnormalities. **Results**: Among the 856 fetuses, 129 fetuses (15.07%) with single positive USM and 36 fetuses (4.21%) with two or more positive USM. There were 81 fetuses (9.46%) with chromosomal abnormalities. In the group with multiple USM, chromosomal abnormalities were found in 36.11% of them. It was higher than the group without USM, which was 6.22% (P < 0.01), and the group with just a single USM (19.38%, P < 0.05). The sensitivity, specificity and accuracy were 96.72%, 98.45% and 98.29% when the combination of USM, Down's syndrome screening and NIPT was used to diagnose fetal chromosomal abnormalities further evaluating the accuracy and effectiveness of the above diagnostic criteria and methods with mainstream Classifiers based evaluation indicators of accuracy, f1 score, AUC. **Conclusions**: The combination of USM, Down's syndrome screening and NIPT is valuable for the diagnosis of fetal chromosomal abnormalities.

Keywords: USM; NIPT; chromosome; fetus; machine learning; data mining

1. Introduction

Non-invasive prenatal testing (NIPT) is transforming prenatal diagnostic practice globally, particularly considering the trend toward late marriage and childbirth while maintaining high sensitivity and specificity for screening for common aneuploidy in the older mother age range [1,2]. Consequently, the likelihood of invasive procedures and the related risk of miscarriage are decreased. NIPT has demonstrated its ability to screen for and diagnose various chromosomal and genetic abnormalities in recent years [2,3]. In recent years. Advanced laboratory testing and purification techniques have improved the performance of NIPT and allowed the introduction of new applications [4,5]. As a consequence of advancements in artificial intelligence, statistical machine learning can ensure correctness and outperform manual efficiency in the analysis and judgment of NIPT findings [6]. With the introduction of this method into clinical practice, NIPT offers tremendous promise for aneuploidy screening [7,8].

However, fewer than half of nations in Europe have included the NIPT test in their national policy, and the test is still utilized by less than 25% of women [9,10]. The thresholds offered to vary considerably in the countries/regions that offer NIPT testing. Belgium and the Netherlands are presently the only two European nations that provide NIPT to all pregnant women [11]. Certain nations provide screening for high-risk pregnancies (e.g., between 1:100 and 1:1000), whereas others limit screening to the highest-risk pregnancies (e.g., above 1:50 or 1:100) [12,13]. In the United States, the majority of insurance companies cover high-risk populations for testing. Additionally, coverage of NIPT chromosomes varies by country/region. Belgium and the Netherlands, in particular, both provide whole genome testing. Belgium is the only nation in the world with a NIPT testing rate of more than 75% [13–15].

Although, the kind and breadth of NIPT tests available and whether the test is publicly supported or must be privately paid vary significantly among health care systems worldwide [3,15,16]. In recent years, metropolitan governments and health care professionals in some developing nations have progressively extended the audience for NIPT to minimize the danger of amniocentesis in women with late pregnancies [4,5]. As the scope of diagnosis broadens, it becomes essential and critical to providing new impetus to the diagnostic process with the help of AI technology. These technologies are known to save exponentially on manual labour, allowing clinicians to concentrate more on verifying unpleasant nonsense occurrences and devising treatment strategies [17,18]. With the rapid development of the Internet of Medical Things, these findings could also serve public health health services well [19–21]. Past reviews of the use of AI in pregnancy research and healthcare have described great potential, but little has been realized in the clinical setting. The low adoption of AI in clinical care in the field of pregnancy may be due to the still unresolved issue of liability for medical errors, especially for so-called black box algorithms. More research on maternal health is needed, with only 31% of papers in the antenatal phase focusing on maternal needs. In recent years, AI techniques have started to gain attention, and supervised learning-based machine learning is naturally applicable to imaging analysis and large datasets in biomedicine. A fusion of various supervised learning may be valuable, as labelled data from supervised ML methods is both time consuming and expensive. Artificial intelligence and machine learning methods can be successfully used to optimize pregnancy outcomes; with appropriate algorithmic improvements, refinements and ethical outcomes, these methods can be incorporated into clinical care [22].

Aki Koivu et al. evaluated the performance of early pregnancy screening for Down's syndrome us-

ing seven predictive machine learning algorithms capable of binary classification, and they concluded that machine learning algorithms could be an adaptive alternative and hold promise for developing better risk assessment models based on existing clinical variables [22]. In a review, Lena Davidson et al. analyze several ML techniques that have been applied to improve chromosome screening, including artificial neural networks (ANN); support vector machines; k-nearest neighbors and deep neural networks, RF, NB, DT and logistic regression (LR). They believe that with appropriate algorithmic improvements, refinement and ethical outcomes, these methods can be incorporated into clinical care [23]. Jianfeng Yang et al. Multiple Z tests using machine learning algorithms can better predict NIPT data. Combining multiple Z values with clinical signs and quality control indicators, features were collected from known samples and scaled for model training in support vector machine (SVM) discriminations, with good results [24]. H.-G. Zhang et al. model and analyze the use of support vector machine algorithms, categorical regression tree algorithms and AdaBoost algorithms in machine learning for prenatal Down's syndrome screening and discuss the feasibility of artificial intelligence algorithms for Down's syndrome screening assessment [25]. Similar studies have been conducted by F. He et al. They conducted separate retrospective studies based on large data case samples and obtained encouraging results by evaluating the training effects of various models, and they still concluded that ML models show robust performance and good extrapolation and can be used as an alternative tool for Down's risk assessment in maternal serum during mi trimester [26–29].

This work was a retrospective study of the treatment course of 856 women diagnosed with high-risk singleton pregnancies in Shenzhen. Their ultrasound (856 instances) and NIPT (642 cases) data from early and mid-pregnancy, respectively, were computed and evaluated using statistical machine learning. Amniocentesis and karyotyping were performed on all 856 individuals to evaluate the efficacy of USM, Down's syndrome screening, and NIPT in identifying fetal chromosomal abnormalities. Additionally, evaluation metrics will be utilized to assess the validity and relevance of AI techniques based on gold standard results. Our approach demonstrates that the addition of NIPT to the USM and Down's syndrome screening process significantly reduces the fact that false positives are high. In terms of accuracy, the mainstream machine learning classifier models, represented by Random Forest, were trained on USM data, Down's syndrome data and NIPT data. The results show that their sensitivity, specificity and accuracy are 96.72%, 98.45% and 98.29%, respectively. Our evaluation metrics demonstrate the feasibility of adding AI technology to the clinical process and provide a realistic foundation for future clinical practice.

The organization of the article is as follows. We presented the USM and NIPT data samples and the data collecting methods utilized in this research in Section 2. Section 3 discusses data processing techniques that correspond to the data structure. Section 4 analyzes the relevant results of several statistical machine learning algorithms. In Section 5, we evaluated and compared clinical study findings. Section 6 concludes the paper.

2. Data and materials

The materials involved in the analysis were clinical data from patients with pregnant women at high risk of having a singleton pregnancy. The dataset was established in Shenzhen Second People's Hospital for the purpose of developing computer-aided diagnosis tools in retrospective pilot studies. All procedures performed in studies involving human participants were under the ethical standards of

the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The dataset's details are given below.

2.1. Study population

According to professional standards guidelines, this is a retrospective study of 856 patients diagnosed as high-risk pregnant women with a singleton pregnancy. They were treated in the Department of Obstetrics and Gynecology of the hospital between January 2017 and December 2019. They were singleton pregnant women. All patients underwent down's screening and karyotype analysis of amniotic fluid to obtain a final diagnosis. Thus, the exclusion criteria were:

- Positive USM, including nuchal translucency (NT) thickening, nuchal fold(NF) thickening, absent or hypoplastic nasal bones, dilated lateral ventricles, dilated renal pelvis, single umbilical artery,intracardiac strong echogenicfoci, and enhanced intestinal echogenicity.
- High risk of serological screening.
- Pregnancy with a history of having a child with chromosomal abnormalities.
- Pregnancy in which one of the spouses is a carrier or inversion of chromosomal balance.
- Pregnancy in which one of the spouses has a fragile X chromosome.
- Pregnancy in which one of the spouses is a patient or carrier of an X linked genetic disorder.
- Pregnancy with a history of the unexplained malformed fetus, spontaneous abortion, stillbirth or neonatal death.
- Pregnancy with one or both spouses having a single gene disease or a history of giving birth to a child with a single gene disease.

All the included patients signed the written informed consent of this study. Among the 856 cases collected, 642 cases received NIPT test, and the remaining 214 cases only received screening and ultrasound examination. According to the inclusion criteria, all the cases studied had high pregnancy risk, and their age and pregnancy information is shown in Table 1.

Table 1. Age and pregnant time distribution of patients with high pregnancy risk.

Variable	Value/Count
Age(mean),range(mean)	20~41 (0.87±5.69)
gestational age(week),range(mean)	12~22 (14.78±3.17)

2.2. Standard methods for clinical testing

2.2.1. USM measurement

Of 856 pregnancies, 642 underwent peripheral blood prenatal NIPT testing as follows: 5 mL of peripheral blood was collected from enrolled pregnant women using EDTA anticoagulation tubes, 32 mixed thoroughly, and plasma was separated at 4°C and stored at -80°C. Plasma cf DNA was extracted with a DNA extraction kit (Shenzhen UW Genetics). A purified library was constructed using a fetal chromosome aneuploidy detection kit, which uses a combined probe anchored ligation sequencing method. Massively parallel sequencing of library DNA was performed using the UW BGISEQ-500 sequencing platform. The human reference genome was applied as a comparator to which the raw data Three obtained from sequencing were compared, and bioinformatics analysis was performed.

2.2.2. Peripheral blood prenatal NIPT testing

All 856 cases underwent amniocentesis at Shenzhen Second People's Hospital, and fetal chromosomes obtained from amniotic fluid were subjected to microarray and karyotype analysis. Traditional cellular chromosome G ribbon karyotype analysis was applied. After centrifugation of the amniotic fluid specimens, they were incubated, fixed, dripped, filmed, stained, and banded, and then read under a microscope by a professional.

2.2.3. Laboratory methods for interventional prenatal diagnosis

At Shenzhen Second People's Hospital, all 856 patients underwent amniocentesis, and fetal chromosomes extracted from amniotic fluid were submitted to microarray and karyotype analysis.

Traditional cellular chromosome G ribbon karyotype analysis was applied. After centrifugation of the amniotic fluid, specimens were incubated, fixed, dripped, filmed, stained, and banded, and then read under a microscope by a professional.

3. Method

3.1. Experimental environment

Hardware environment: CPU 4 cores, RAM 32GB, GPU v100, video memory 16GB, disk 100GB Environment.

Configuration: Python version 3.7, framework version Sklearn.
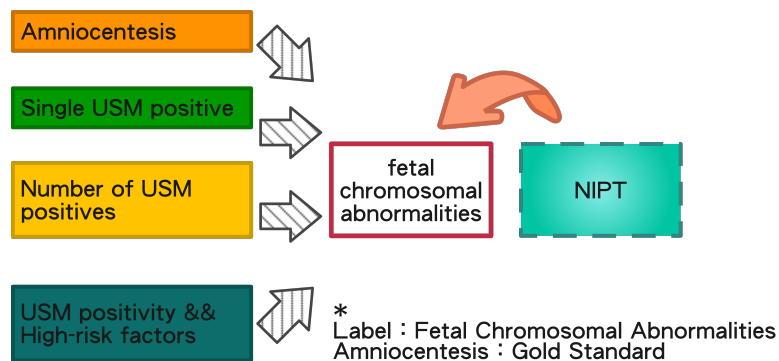
3.2. The design of the experiment

Figure 1. Schematic diagram of the experimental design process.

3.2.1. Experimental procedure

The experiment was built on the assumption that data is the artificial intelligence ceiling. As indicated in Figure 1, the true values of the data set of the investigation were first statistically and analytically evaluated using SPSS 22.0. The relationship between a) amniocentesis and chromosomal findings, b) single USM positivity and chromosomal abnormalities, c) number of USM positivity and chromosomal abnormalities, d) diagnostic efficacy of USM positivity and combined high-risk factors

for fetal chromosomal abnormalities and, e) relationship between NIPT test and amniocentesis kary-otype results were analyzed computationally, respectively. The effectiveness of adding NIPT to clinical diagnosis is compared and to utilized to evaluate the AI technology's maximal effectiveness. Secondly, we used the USM and Down's syndrome screening data alone as control group datasets and used statistical machine learning algorithms for computation and training. Furthermore, the control group results were compared and analyzed with the true values to obtain the feasibility of implementing NIPT as a supplementary diagnostic tool with AI technology.
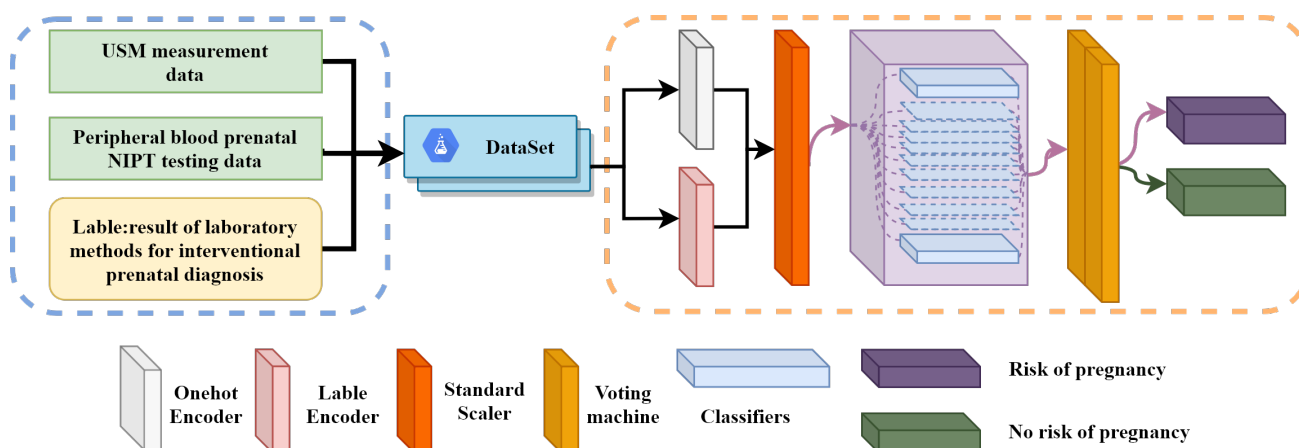

Figure 2. Experimental flowchart.

3.2.2. Statistical methods and preset parameters

SPSS22.0 software was applied to statistically analyze the data. Count data were expressed as frequency or rate, test or Fisher's exact probability method was used for comparison between groups, and ROC curve method was used for diagnostic efficacy analysis; differences were considered statistically significant at $P < 0.05$.

3.2.3. Data preprocessing

To ensure the accuracy of the operation, first set all the features of all numeric types to single-4 precision floating-point types. A key assumption of all verification methods is the independence of test set and train set.

All validation methods rely on separating the dataset into training and testing (and sometimes validation) subsets, i.e., the data are trained on the training subset and tested on the testing subset, which should be unavailable to the method that is being validated during the training process. In this study, we divided the data set into the training set and the test set according to the ratio of 70% and 30%, and the test set was only used in the evaluation stage. If there are missing values in the sample, it may adversely affect the calculation results. Therefore, we checked the completeness of all samples to ensure that there were no missing values. The distribution of data labels is very important for model calculations. We performed one-hot encoding on the feature 14 labels and encoded the result through label-encoder (Figure 1).

As we all know, most medical samples are unbalanced, and this research is still the case(Figure 2).
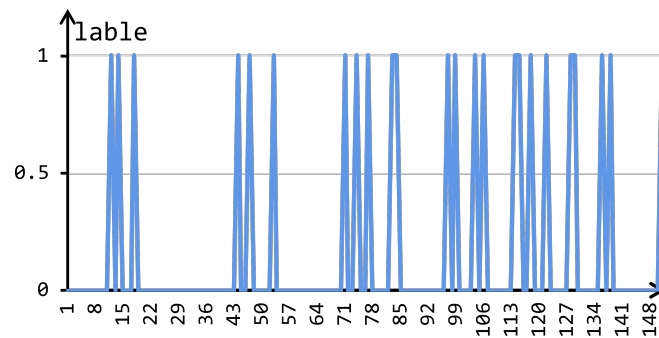
Figure 3. Proportional distribution of high-risk or no-risk pregnant women. (1:high-risk 0:no-risk).
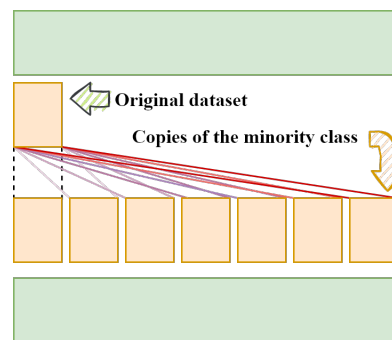

Figure 4. A schematic depicting SMOTE subsampling method logic implemented.

Therefore, we have up sampled the samples before training the model each time. Figure 3 introduces the SMOTE up sampling idea used in this study to analyze and simulate a few types of samples and add new artificially manufactured samples to the data set. In turn, the categories in the original data are no longer seriously unbalanced. This is to ensure that the learned data is balanced during the training processes of the model.

3.2.4. Classifier settings

This research mainly uses Random forest, logistic regression and support vector machine(SVM) to fit the samples in the training set.

3.2.5. Parameter settings

In Random Forest, the function for measuring segmentation quality is set as the Gini function. However we do not set the maximum depth of the tree, which means that during the fitting process, the model will be extended to all leaves are pure.

In the logistic regression classifier, as an optimization problem, the binary class l2 penalty logistic regression minimizes the following cost function:

$$\min_{w,c} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \log \left(\exp \left(-y_i \left(X_i^T w + c \right) \right) + 1 \right),$$ (3.1)

Similarly, l1 regularized logistic regression solves the following optimization problem:

$$\min_{w,c} \|w\|_1 + C \sum_{i=1}^{n} \log \left(\exp \left(-y_i \left(X_i^T w + c \right) \right) + 1 \right), \tag{3.2}$$

Elastic-Net regularization is a combination of ℓ_1 and ℓ_2, and minimizes the following cost function:

$$\min_{w,c} \frac{1-\rho}{2} w^T w + \rho \|w\|_1 + C \sum_{i=1}^{n} \log \left(\exp \left(-y_i \left(X_i^T w + c \right) \right) + 1 \right), \tag{3.3}$$

where ρ controls the strength of ℓ_1 regularization vs. ℓ_2 regularization.

In the support vector machine, we set RBF as the kernel function. Among them, γ is determined by the parameter "gamma" and must be greater than 0.

To evaluate the predictive models used, a 10-fold cross-validation process was carried out. K-fold cross-validation is an evaluation tool that, in this study, tests the accuracy and ability to generalize to new data by dividing the used dataset into 10 subsets and performing the model training and evaluation phases 10 times. The advantage of using k-fold cross-validation is that the predictions are less sensitive to data partitioning as the data are partitioned into different training and test sets k times. The 10-fold cross-validation was chosen to fit all quantities in the training set. It is worth noting that even though k-fold cross-training is less sensitive to data partitioning, we still perform cross-training on the training set. In other words, the data in the test set is not seen by the model during the training phase, which ensures the rigour with which we evaluate the accuracy and generalization ability of the model on the test set. K-fold cross-validation maintained the parameter settings of all models, and to control for variables, we did not adjust the parameters of all classification models during their initialization, keeping only the simplest configuration rather than complicating their configuration. At the end of training, all classifier models are used to classify the test dataset. The classification results will be compared based on the set evaluation metrics.

3.2.6. Evaluating markers

Given the abundance of the proposed methods for detecting ischemic beats and MI, it is essential to validate these methods and compare them to find the more effective ones. Therefore, it's critical to utilize reliable validation methods and consistent performance measurements to provide generalizable and repeatable outcomes.

First, there is a lack of uniformity in the performance measures that are used for evaluation. Most of the studies use classification accuracy, which is a common measure of performance in technical fields. The classification performance is mainly evaluated by the classification accuracy, which is defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{3.4}$$

where TP, TN, FP and FN are the true positive, the true negative, the false positive and the false negative, respectively. Besides the classification accuracy, the other two measurements, F1-score, ROC and AUC, are used to evaluate classification performance. F1-score is defined as the harmonic mean of precision and recall.

$$F_1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \tag{3.5}$$

where

$$\text{precision} = \frac{TP}{TP + FP}, \tag{3.6}$$

and

$$\text{recall} = \frac{TP}{TP + FN}. \tag{3.7}$$

4. Results

4.1. Results of amniocentesis and chromosomal examination

After amniocentesis and chromosome examination in 856 pregnancy, a total of 81 cases of chromosomal abnormalities and 775 cases of normal chromosomes were detected.

4.2. Relationship between single USM positivity and chromosomal abnormalities

There were 129 cases (15.07%) of single USM positivity, of which 25 patients (19.38%) had chromosomal abnormalities. Among the solitary USM positivity, NT thickening and nasal bone abnormalities had the greatest diagnostic value for fetal chromosomal abnormalities (P < 0.05), as shown in Table 2.

Table 2. Relationship between solitary USM positivity and chromosomal abnormalities.

USM positivity		Chromosomal abnormalities (cases)		χ^2 value	P value
		Yes	None		
NT thickening	Yes	14	67	6.388	0.011
	None	67	708		
Nasal bone abnormalities	Yes	5	15	4.063	0.044
	None	76	760		
NF thickening	Yes	1	3	-	0.329
	None	80	772		
Widening of the lateral ventricles	Yes	1	3	-	0.329
	None	80	772		
Dilated renal pelvis	Yes	0	4	-	1.000
	None	81	771		
Single umbilical artery	Yes	2	5	-	0.136
	None	79	770		
Intracardiac strong echogenic foci	Yes	0	2	-	1.000
	None	81	773		
Echo enhancement of the intestinal canal	Yes	1	2	-	0.258
	None	80	773		
Choroid plexus cyst	Yes	1	3	-	0.329
	None	80	772		

4.3. Relationship between the number of USM positives and chromosomal abnormalities

The rate of chromosomal abnormalities in the USM positive number 1 group was 19.38%, which was higher than that in the USM negative group (6.22%) (χ^2 =24.744, P < 0.01); the rate of chromosomal abnormalities in the USM positive number 2 group was 36.11%, which was higher than that in the USM negative group (6.22%) (χ^2 = 42.994, P < 0.01), and also higher than that in the USM positive number 1 group 19.38% (χ^2 = 4.445, P < 0.05), as shown in Table 3.

Table 3. Relationship between solitary USM positivity and chromosomal abnormalities.

Number of USM positives	Chromosomal abnormalities (cases)		χ^2 value	P value
	Yes	None		
0	43	648		
1	25*	104	53.116	0.000
≥2	13*#	23		

Note:* indicates that the rate of chromosomal abnormalities in this group has P<0.05 compared with the USM-negative group, and # indicates that the rate of chromosomal abnormalities in this group has P<0.05 compared with the USM-positive number 1 group.

4.4. Diagnostic efficacy of USM positivity and combined high-risk factors for fetal chromosomal abnormalities

The results of amniocentesis karyotyping were used as the diagnostic gold standard. There were 274 (32.01%) positive USM screening cases alone, of which 57 (20.80%) were chromosomal abnormalities. 265 (30.96%) positive USM + Down screening cases, of which 62 (23.40%) were chromosomal abnormalities. 68 (10.59%) positive USM + Down screening + NIPT screening cases, of which 59 (86.76%) were chromosomal abnormalities. The sensitivity, specificity and accuracy of USM+ Down screening + NIPT screening for the diagnosis of fetal chromosomal abnormalities were 96.72% and 98.45% and 98.29% respectively, as shown in Table 4.

Table 4. Diagnostic efficacy of USM positivity and combined high-risks factors for fetal chromosomal abnormalities.

Screening index		Chromosomal abnormalities (cases)		Sensitivity (%)	Specificity (%)	Accuracy (%)	P predictive value (%)	N predictive value (%)
		Yes	None					
USM alone	P	57	217	70.37	72.00	71.85	20.80	95.88
	N	24	558					
USM+Down screening	P	62	203	76.54	73.81	74.07	23.40	96.79
	N	19	572					
USM+Down screening+NIPT	P	59	9	96.72	98.45	98.29	86.76	99.65
	N	2	572					

Note: "P" means Positive and "N" means Negative.

4.5. Comparison of NIPT test and amniocentesis karyotype results

In this study, 68 patients with "USM positive + Down screening + NIPT positive" , NIPT suggested 34 patients with a high risk of trisomy 21, 23 patients with an increased risk of trisomy 21 and trisomy 18, and 11 patients with an increased risk of trisomy 13, 86.76% (59/68) were consistent with the karyotype results. The mean age was (34.67 ± 4.77) years, of which 36 patients were >35 years.

4.6. Evaluation results of the classifier

The evaluating indicators among USM based datasets are significantly different from those of USM combined with NIPT in several key respects. Figure 5 displayed the four mainly evaluating indicators of train set and test set of USM data.
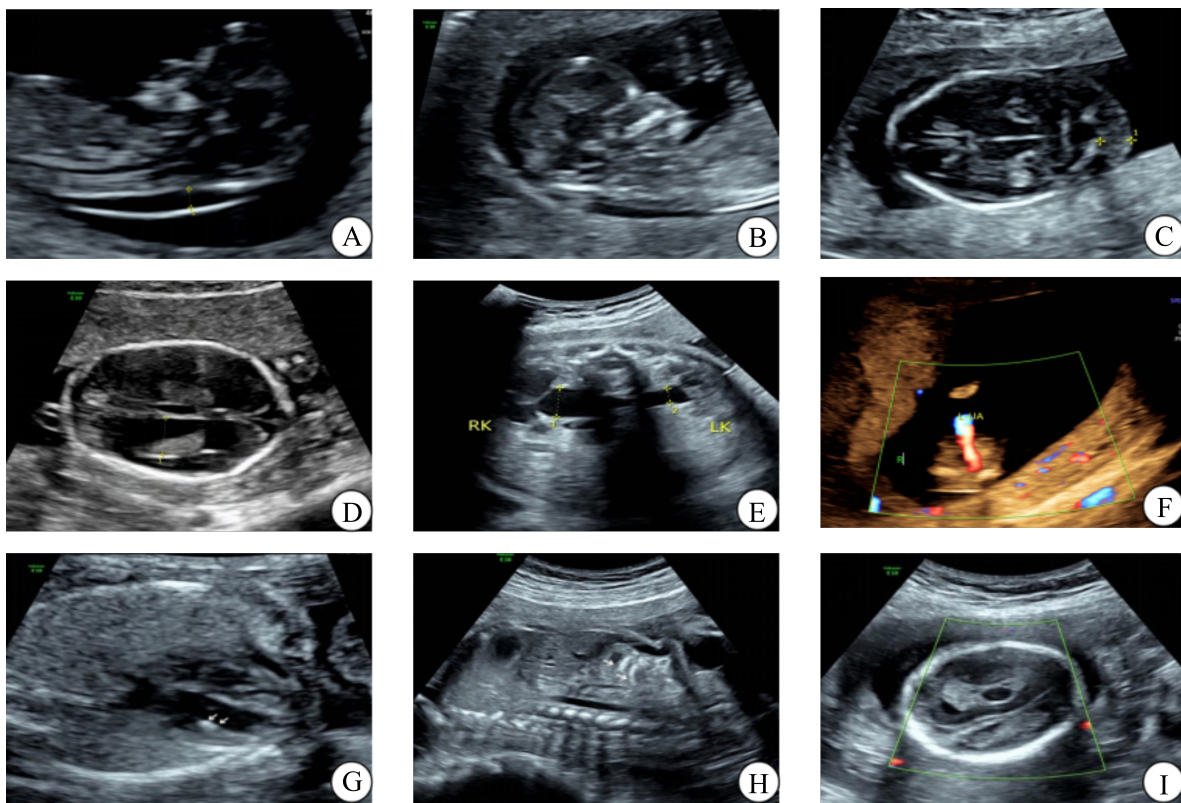

Figure 5. Positive sonogram of various ultrasonic soft indicators.

For the control variables, we configured the same parameters for each classifier and did not set additional parameters. In terms of the presentation of the results, there were significant differences, with the Random Forest, KNN, Bagging, Gradient Boosting, and Decision Tree classifiers showing good performance in terms of ACC and f1 score, but the MLP, Extra tree The performance of MLP, Extra tree and Decision Tree classifier is good in terms of ACC and f1 score.

We especially found that AUC exhibits significant disparities of classifiers involved in figure(D). This proves inconsistent results for prediction results and the value for approach based NIPT techniques.
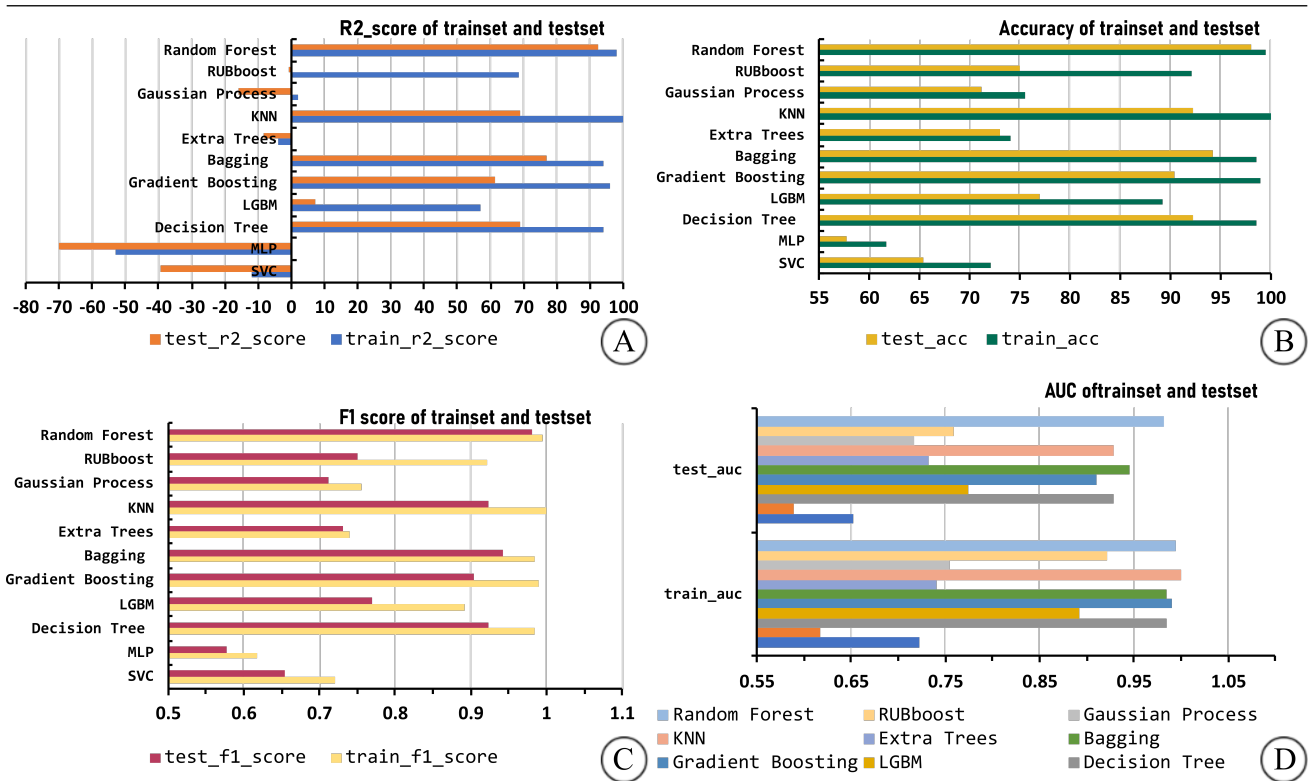
Figure 6. Evaluation results of classifier on train set and test set.
(A:the r2 score of train set and test set B:Accuracy of train set and test set C:F1 score of train set and test set D:AUC of train set and test set)

Overall, non-NIPT types of data were learned on different classifiers with poor results, despite some data pre-processing. This reflects the value of the importance of NIPT data.

5. Discussion

Certain specific signs of prenatal ultrasound suggest the possibility of chromosomal abnormalities in the fetus, and their positive presentation may suggest further invasive prenatal testing such as amniocentesis [30]. The results of this study showed that the detection rate of these indices did not agree with the diagnostic value: ① NT thickening, recognized as one of the most predictive USM, was associated with fetal abnormalities such as multiple organ development or chromosomal abnormalities. In this study, 81 cases of NT thickening had the highest incidence of all USMs. Among them, 14 cases (17.28%) had chromosomal abnormalities, which is like foreign studies [31],and is of great significance for the early detections of fetal chromosomal abnormalities. ② Abnormal nasal bone development, an important USM indicator. In the present study, 5 out of 20 cases (25.00%) of nasal bone abnormalities were chromosomal abnormalities, similar to those reported in the literature [32]. The detection of nasal bone abnormalities is difficult in ultrasonography and should prevent missed misdiagnosis. ③ There is a correlation between NF thickening, and chromosomal aberrations. There were only four cases of NF thickening in our group, and one of them had chromosomal abnormalities, which may be related to the composition of our circumstances. No further NF examination was per-

formed in the cases with no positive findings on NT examination in our group, resulting in fewer cases with abnormal NF. ④ Other USM positivity, including lateral ventricular dilatation, renal pelvic dilatation, single umbilical artery, intracardiac strong echogenic foci, intestinal echogenic enhancement, and choroidal cysts have some diagnostic value for fetal chromosomal abnormalities, particularly in fetuses with multiple USM positivity.

In the present study, all the above-mentioned USM-positive cases were less frequent: one case with ultrasound suggestive of several strong spots in the left and right ventricles, a right lateral ventricular choroid plexus cyst, and amniocentesis showing a de novo mutated micro repeat in the long arm of chromosome 22; one case with enhanced bowel echo combined with fetal 4th ventricular widening and right ectopic kidney, whose amniocentesis showed normal karyotype but positive for rubella virus infection; In one case of NT thickening combined with choroid plexus cyst amniocentesis did not reveal chromosomal abnormalities and the choroid plexus cyst disappeared on repeat ultrasound at 26 weeks of gestation.

The literature [33] reported an increased risk of fetal trisomy 18 in those with choroid plexus cysts that did not disappear by 26 weeks of gestation. The study [34] showed that the predictive accuracy of chromosomal abnormalities with multiple USMs was significantly better than that with a single USM.

In this study, the risk of chromosomal abnormalities was higher in fetuses with multiple ($n \geq 2$) positive USMC than in fetuses with single ($n \geq 1$) positive USMC, implying that the clinical significance of USMC has a cumulative effect and that the more USMC detected on ultrasound screening, the more pregnancy should be advised to undergo further investigation.

NIPT screening has been widely used in clinical practice worldwide [35, 36]. The results of this study found that USM-positive combined with Down's screening and NIPT high-risk improved the sensitivity, specificity and accuracy of the diagnosis of fetal chromosomal abnormalities by 96.72%, 98.45% and 98.29%, respectively, which fully indicates that NIPT screening is necessary for fetuses with USM and Down's screening abnormalities to further clarify the risk of developing fetal chromosomal abnormalities. In the present study, 68 patients with positive USM combined with Down's screening and high risk of NIPT had a high compliance rate (86.76%) with karyotype results, indicating the high accuracy of NIPT testing for the diagnosis of fetal chromosomal abnormalities, like the results of previous studies [37]. The importance of ultrasound screening is also illustrated by the fact that all 59 fetuses with abnormal NIPT and confirmed chromosomal abnormalities by amniocentesis in our group of cases had abnormal ultrasound soft makers.

The main shortcomings of this study are that the number of cases is not large enough, especially the number of cases of USM positive fetuses other than NT thickening and nasal bone abnormalities, to make an accurate determination of the clinical value of these USM abnormalities. All cases enrolled in this study were singleton pregnancies, and there is a lack of data from studies on the twin or multiple pregnancies. In addition, due to the high cost of the NIPT testing, not all fetuses in this study were 12 tested, which may bias the study results.

6. Conclusions

In summary, among pregnancies with high-risk factors, the chance of chromosomal abnormalities in USM-positive fetuses in early and mid-term pregnancies is higher than in USM-negative fetuses. The predictive value of different USM types and the number of USM abnormalities for fetal chromosomal

abnormalities varies. USM combined with Down's screening and NIPT testing is of great value in detecting fetal chromosomal abnormalities. When the prenatal diagnosis is found to have positive fetal USM, a comprehensive and integrated assessment should be performed in combination with the maternal age and serological findings. NIPT testing must be performed on high-risk fetuses, and amniocentesis chromosome examination should be performed if necessary to improve the accuracy of diagnosis. In contrast, unnecessary interventional prenatal diagnosis on low-risk pregnant women should be avoided as much as possible.

On the one hand, statistical machine learning algorithms offer a viable alternative to prenatal risk assessment, with better detection rates at certain false positive rates reducing unnecessary invasive testing, while the introduction of the technology into prenatal diagnosis can effectively reduce hospital screening costs and is suitable for use in large scale populations. Beyond this, the main limitation of our work is the relatively small sample size. In order to fully investigate the capabilities of statistical machine learning algorithms, we need to develop, train and test algorithms using larger datasets, including a wider range of clinical and demographic variables.

Acknowledgments

Conflict of interest

The authors declare there is no conflict of interest.

References

1. M. Allyse, M. A. Minear, E. Berson, S. Sridhar, M. Rote, A. Hung, et al., Non-invasive prenatal testing: a review of international implementation and challenges, *Int. J. Women's Health*, **7** (2015), 113. https://doi.org/10.2147/IJWH.S67124

2. R. W. Chiu, Y. D. Lo, Non-invasive prenatal diagnosis by fetal nucleic acid analysis in maternal plasma: the coming of age, *Semin. Fetal Neonatal Med.*, **16** (2011), 88–93. https://doi.org/10.1016/j.siny.2010.10.003

3. Z. Ren, Z. Chai, Z. Zhou, China fertility report 2019, *The Development research*, **6** (2019), 20–40.

4. L. Yang, L. Hongtian, Z. Yubo, L. Jianmeng, Crowd fertility and fertility protection, *Chin. J. Reprod. Health*, **31** (2020), 401–403. https://doi.org/10.1016/j.cclet.2019.06.048

5. Z. Xuegui, Y. Jiang, X. Guige, Retrospective analysis of pregnancy risk of older parturients, *J. Binzhou Med. Univ.*, **43** (2018), 118–121.

6. E. Hung, R. Chiu, Y. Lo, Detection of circulating fetal nucleic acids: a review of methods and applications, *J. Clin. Pathol.*, **62** (2009), 308–313. https://doi.org/10.1136/jcp.2007.048470

7. G. J. Liao, A. M. Gronowski, Z. Zhao, Non-invasive prenatal testing using cell-free fetal dna in maternal circulation, *Clin. Chim. Acta*, **428** (2014), 44–50. https://doi.org/10.1016/j.cca.2013.10.007

8. K. H. Nicolaides, A. Syngelaki, G. Ashoor, C. Birdir, G. Touzet, Noninvasive prenatal testing for fetal trisomies in a routinely screened first-trimester population, *Am. J. Obstet. Gynecol.*, **207** (2012), 374.e1–6. https://doi.org/10.1016/j.ajog.2012.08.033

9. D. W. Bianchi, L. D. Platt, J. D. Goldberg, A. Z. Abuhamad, A. J. Sehnert, R. P. Rava, Genome-wide fetal aneuploidy detection by maternal plasma dna sequencing, *Obstet. Gynecol.*, **119** (2012), 890–901. https://doi.org/10.1097/AOG.0b013e31824fb482

10. G. E. Palomaki, E. M. Kloza, G. M. Lambert-Messerlian, J. E. Haddow, L. M. Neveux, M. Ehrich, et al., Dna sequencing of maternal plasma to detect down syndrome: an international clinical validation study, *Genet. Med.*, **13** (2011), 913–920. https://doi.org/10.1097/GIM.0b013e3182368a0e

11. K. Song, T. J. Musci, A. B. Caughey, Clinical utility and cost of non-invasive prenatal testing with cfdna analysis in high-risk women based on a us population, *J. Matern.-Fetal Neonatal Med.*, **26** (2013), 1180–1185. https://doi.org/10.3109/14767058.2013.770464

12. M. Allyse, L. C. Sayres, T. A. Goodspeed, M. K. Cho, Attitudes towards non-invasive prenatal testing for aneuploidy among us adults of reproductive age, *J. Perinatol.*, **34** (2014), 429–434. https://doi.org/10.1038/jp.2014.30

13. L. S. Chitty, D. W. Bianchi, Noninvasive prenatal testing: the paradigm is shifting rapidly, 2013. https://doi.org/10.1002/pd.4136

14. Y. D. Lo, K. A. Chan, H. Sun, E. Z. Chen, P. Jiang, F. M. Lun, et al., Maternal plasma dna sequencing reveals the genome-wide genetic and mutational profile of the fetus, *Sci. Transl. Med.*, **2** (2010), 61ra91–61ra91. https://doi.org/10.1126/scitranslmed.3001720

15. S. Morris, S. Karlsen, N. Chung, M. Hill, L. S. Chitty, Model-based analysis of costs and outcomes of non-invasive prenatal testing for down's syndrome using cell free fetal dna in the uk national health service, *PloS One*, **9** (2014), e93559. https://doi.org/10.1371/journal.pone.0093559

16. A. C. Neocleous, K. H. Nicolaides, C. N. Schizas, First trimester noninvasive prenatal diagnosis: A computational intelligence approach.

17. WANGhui, Pregnancy risk analysis for older parturients, *Chin. Commun. Physician*, **19**, 2019.

18. G. ZHAN, Study on risk assessment and prevention and control strategies of adverse pregnancy outcomes for women in the third trimester, Ph.D. dissertation, Changchun: Jilin University, 2019.

19. W. Wu, S. Pirbhulal, G. Li, Adaptive computing-based biometric security for intelligent medical applications, *Neural Comput. Appl.*, **32** (2020), 11 055–11 064.

20. A. H. Sodhro, M. S. Al-Rakhami, L. Wang, H. Magsi, N. Zahid, S. Pirbhulal, et al., Decentralized energy efficient model for data transmission in iot-based healthcare system, in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring).* IEEE, 2021, pp. 1–5. https://doi.org/10.1109/VTC2021-Spring51267.2021.9448886

21. A. H. Sodhro, A. Gurtov, N. Zahid, S. Pirbhulal, L. Wang, M. M. U. Rahman, et al., Toward convergence of ai and iot for energy-efficient communication in smart homes, *IEEE Internet Things*, **8** (2020), 9664–9671. https://doi.org/10.1109/JIOT.2020.3023667

22. A. Koivu, T. Korpimäki, P. Kivelä, T. Pahikkala, M. Sairanen, Evaluation of machine learning algorithms for improved risk assessment for down's syndrome, *Comput. Biol. Med.*, **98** (2018), 1–7. https://doi.org/10.1016/j.compbiomed.2018.05.004

23. L. Davidson, M. R. Boland, Towards deep phenotyping pregnancy: a systematic review on artificial intelligence and machine learning methods to improve pregnancy outcomes, *Brief. Bioinformatics*, 2021. https://doi.org/10.1093/bib/bbaa369

24. J. Yang, X. Ding, W. Zhu, Improving the calling of non-invasive prenatal testing on 13-/18-/21-trisomy by support vector machine discrimination, *BioRxiv*, p. 216689, 2017. https://doi.org/10.1101/216689

25. H.-G. Zhang, Y.-T. Jiang, S.-D. Dai, L. Li, X.-N. Hu, R.-Z. Liu, Application of intelligent algorithms in down syndrome screening during second trimester pregnancy, *World J. Clin. Cases*, **9** (2021), 4573. https://doi.org/10.12998/wjcc.v9.i18.4573

26. F. He, B. Lin, K. Mou, L. Jin, J. Liu, A machine learning model for the prediction of down syndrome in second trimester antenatal screening, *Clinica Chimica Acta*, **521** (2021), 206–211. https://doi.org/10.1016/j.cca.2021.07.015

27. Ö. Uzun, H. Kaya, F. Gürgen, and F. G. Varol, Prenatal risk assessment of trisomy 21 by probabilistic classifiers, in *2013 21st Signal Process. Commun. Appl. Conf. (SIU)*. IEEE, 2013, pp. 1–4. https://doi.org/10.1109/SIU.2013.6531604

28. H. Ocak, A medical decision support system based on support vector machines and the genetic algorithm for the evaluation of fetal well-being, *J. Med. Syst.*, **37** (2013), 1–9. https://doi.org/10.1007/s10916-012-9913-4

29. I. Vovsha, A. Rajan, A. Salleb-Aouissi, A. Raja, A. Radeva, H. Diab, et al., Predicting preterm birth is not elusive: Machine learning paves the way to individual wellness, in *2014 AAAI Spring Symposium Series*, 2014.

30. L. Zhen, Y.-D. Yang, Y.-J. Li, L.-L. Xu, D.-Z. Li, The role of ultrasound in the choice between chorionic villus sampling and amniocentesis for patients with a positive nipt result for trisomy 18/13, *Prenat. Diagn.*, **39** (2019), 1155–1158. https://doi.org/10.1002/pd.5524

31. S. K. Alldred, Y. Takwoingi, B. Guo, M. Pennant, J. J. Deeks, J. P. Neilson, et al., First trimester ultrasound tests alone or in combination with first trimester serum tests for down's syndrome screening, *Cochrane Database Syst. Rev.*, no. 3, 2017. https://doi.org/10.1002/14651858.CD012600

32. K. M. Wojda, H. Moczulska, P. J. Sieroszewski, The absence of fetal nasal bones in ultrasound examination between 11+ 0 and 13+ 6 weeks of gestation versus the occurrence of trisomies 21, 18, and 13, *Ginekol. Pol.*, **90** (2019), 604–606. https://doi.org/10.5603/GP.2019.0104

33. Y. Nan, Z. Liu, J. Zhang, R. Yang, Q. Yu, F. Wang, et al., Relationship between the choroid plexus cyst and the aneuploidy, *Zhonghua Yi Xue Za Zhi*, **98** (2018), 2987–2990.

34. A. Åhman, O. Axelsson, G. Maras, C. Rubertsson, A. Sarkadi, P. Lindgren, Ultrasonographic fetal soft markers in a low-risk population: prevalence, association with trisomies and invasive tests, *Acta Obstet. Gynecol. Scand.*, **93** (2014), 367–373. https://doi.org/10.1111/aogs.12334

35. T. S. Hartwig, L. Ambye, L. Werge, M. K. Weiergang, P. Nørgaard, S. Sørensen, et al., Non-invasive prenatal testing (nipt) in pregnancies with trisomy 21, 18 and 13 performed in a public setting–factors of importance for correct interpretation of results, *Eur. J. Obstet. Gynecol. Reprod. Biol.*, **226** (2018), 35–39. https://doi.org/10.1016/j.ejogrb.2018.04.042

36. J. Harraway, Non-invasive prenatal testing, *Aust. Fam. Physician*, **46** (2017),735–739.

37. A. J. Ainsworth, M. A. Holman, E. Codsi,M. Wick, Use of genetic testing after abnormal screening ultrasound: A descriptive cohort study, *Gynecol. Obstet. Invest.*, **83** (2018), 466–470. https://doi.org/10.1159/000484242

AIMS Press