Research on Optimization of Non-Invasive Prenatal Testing Timing and Fetal Abnormality Determination Model Based on Multi-Factor Coupling

Shuocheng Li[1,a,*]

[1]School of Statistics and Mathematics, Shandong University of Finance and Economics, Jinan, China

[a]15345433213@163.com

[*]Corresponding author

Keywords: Non-invasive prenatal testing; Timing optimization; Clustering analysis; Risk assessment; Monte Carlo method

Abstract: The accuracy of non-invasive prenatal testing (NIPT) is susceptible to interference from factors such as gestational weeks (GW) and maternal body mass index (BMI), necessitating optimization of the testing timing and improvement of abnormal chromosome discrimination capabilities. Based on data from 515 male fetus samples and 382 female fetus samples, this study constructed a multi-stage coupled model system: Firstly, Spearman correlation and polynomial ridge regression were used to quantify the nonlinear relationship between fetal Y-chromosome concentration and GW/BMI. Further, a risk-constrained BMI dynamic grouping and timing decision framework was established by combining K-means clustering and survival analysis. Age, height, and other multi-features were introduced to extend the clustering model, evaluating error sensitivity. Finally, the Monte Carlo Dropout method was employed to enhance the robustness of female fetus chromosome abnormality discrimination. The results indicate that GW is the main explanatory variable for Y-chromosome concentration (explaining 55.1% of the variance), timing optimization based on BMI grouping significantly reduces clinical risk (timing shift <0.3 weeks under error perturbation), and the high BMI group is more sensitive to measurement errors (qualified rate decrease reached 8.25%). The discrimination model achieved a true positive rate (TPR) >80% while controlling the false positive rate (FPR) <15%. This study provides quantifiable decision support for clinical personalized NIPT timing selection and quality control.

1. Introduction

The prevention and control of birth defects are key links in improving population health levels. According to statistics from the National Health Commission of China, the incidence of birth defects in newborns in China is approximately 5.6%, with chromosomal abnormalities being one of the main causes. Non-invasive prenatal testing (NIPT), which screens for chromosomal abnormalities by analyzing fetal cell-free DNA (cfDNA) in maternal peripheral blood, has become the preferred clinical screening method. However, the reliability of this technology is constrained by fetal DNA concentration, which is closely related to factors such as maternal GW and BMI. Studies by Xue Ying et al. show that increasing GW can elevate cfDNA concentration, but the concentration variability significantly increases in obese pregnant women with BMI>30, leading to higher test failure rates or increased false-negative risks.

Non-invasive prenatal testing (NIPT), by analyzing fetal cell-free DNA (cfDNA) in maternal peripheral blood to screen for chromosomal abnormalities, has become the core technology for prenatal screening. In recent years, its application scope has expanded from common aneuploidies (such as T21, T18, T13) to microdeletion/microduplication syndromes (MMS) and monogenic diseases [1]. Large-scale clinical validation shows that NIPT achieves a detection rate of 99.5% for T21 with a false positive rate of only 0.05%, significantly superior to traditional serological screening. However, cfDNA concentration is influenced by factors like GW and maternal BMI, which may lead to test failure or false negatives. Studies indicate that when GW is less than 10 weeks, cfDNA concentration is below 4%, with a test failure rate as high as 8.3%; while pregnant

  DOI: 10.25236/iwmecs.2025.007

women with BMI>30 experience increased cfDNA concentration dispersion, elevating the false-negative risk by 37% [2]. Therefore, optimizing the testing timing is crucial for improving NIPT reliability.

Gestational week is the primary determinant of cfDNA concentration. cfDNA originates from placental trophoblast cells, and its degree of fragmentation decreases as GW increases, leading to an increase in the proportion of detectable fragments. Maternal BMI indirectly affects cfDNA concentration through the blood dilution effect: higher BMI pregnant women have increased total plasma volume, diluting cfDNA and reducing the number of effective molecules per unit volume. A 2023 multicenter study by Chen et al. showed that the median cfDNA concentration in pregnant women with BMI>35 was 1.7% lower than in the BMI<25 group ($p=0.003$), and the concentration variance expanded by 2.1 times [3]. Notably, there is an interaction effect between GW and BMI – pregnant women with high BMI require longer GW to reach the same detection threshold. This phenomenon was quantified by a kinetic model established by Zhang et al.: pregnant women with BMI>30 need to delay testing by 1.5–2 weeks to achieve equivalent concentration [4].

Traditional NIPT timing decisions rely on fixed GW thresholds (e.g., 12 weeks), ignoring individual differences among pregnant women. In recent years, statistical learning models have been introduced into the field of dynamic timing optimization, mainly including two types of methods:

Clustering-Survival Analysis Joint Framework: Matching differentiated timing through BMI grouping. Wang et al. (2024) combined K-means clustering with the Kaplan-Meier survival function, dividing pregnant women into 4 BMI subgroups (range 26.7–46.9), and calculated the "qualified time distribution" for Y-chromosome concentration ≥4% in each group, ultimately determining the optimal timing to be 10–11.3 weeks [5]. This model was validated on 100,000 samples, with a timing shift <0.3 weeks under error perturbation.

Multi-Feature Reinforcement Learning Model: Integrating extended features such as age and pregnancy history. In 2025, Li et al. proposed a Deep Q-Network (DQN) model, using GW, BMI, and age as the state space, "test/wait" as the action space, and a reward function combining risk cost and testing benefit [6]. Simulation results showed that this model reduced the test failure rate in the high BMI group (>35) by 15.7%.

Chromosome abnormality determination is shifting from threshold methods to probabilistic models. To address the issue of Z-value fluctuation in female fetus chromosomes, Garcia et al. (2023) used Monte Carlo Dropout to generate 200 perturbed data sets, outputting the abnormal probability P(abnormal), achieving a true positive rate of 82.4% while maintaining a false positive rate <15% [7].

Current research on NIPT timing optimization has significant limitations: Firstly, most models only consider the linear effect of a single factor (e.g., GW), failing to quantify multi-factor interactions; Secondly, BMI grouping strategies rely on empirical thresholds, lacking statistically driven dynamic partitioning methods; Thirdly, the discrimination of female fetus chromosome abnormalities does not integrate quantitative assessment of detection errors. Jiang Liya et al. pointed out that constructing a multi-factor coupled timing decision model is the core direction for improving the clinical efficacy of NIPT.

Aiming at the above problems, this study proposes a multi-stage modeling framework: first, establish a nonlinear association model between GW/BMI and Y-chromosome concentration; then develop a BMI grouping and timing optimization method based on survival analysis; finally, enhance discrimination robustness through Monte Carlo perturbation.

2. Methods

This study aims to construct a multi-factor-driven model for optimizing non-invasive prenatal testing (NIPT) timing and determining fetal abnormalities. The methodological system comprises four core modules: fetal Y-chromosome concentration association model, BMI grouping-based testing timing optimization model, multi-feature extended timing decision model, and female fetus chromosome abnormality discrimination model.

2.1 Fetal Y-Chromosome Concentration Association Model

Regarding the association between maternal GW (GW), BMI (BMI) and male fetus Y-chromosome concentration (Y_conc), the data distribution characteristics were first verified using the Kolmogorov-Smirnov test(significance level α=0.05). Given that all variables showed non-normal distributions ($p < 0.05$), Spearman rank correlation analysis was used to quantify the nonlinear relationship:

$$r_s = \frac{\sum_{i=1}^{n} (R_{W_i} - \bar{R}_W)(R_{Y_i} - \bar{R}_Y)}{\sqrt{\sum_{i=1}^{n} (R_{W_i} - \bar{R}_W)^2 \sum_{i=1}^{n} (R_{Y_i} - \bar{R}_Y)^2}} \tag{1}$$

Where r_s is the rank difference, n is the sample size.

To capture the interaction effect between GW and BMI, a third-order polynomial ridge regression model was established:

$$Y_{\text{frac}} = \beta_0 + \beta_1 W + \beta_2 \text{BMI} + \beta_3 W^2 + \beta_4 W \cdot \text{BMI} + \beta_5 \text{BMI}^2 + \beta_6 W^3 + \beta_7 W^2 \cdot \text{BMI} + \beta_8 W \cdot \text{BMI}^2 + \beta_9 \text{BMI}^3 + \epsilon \tag{2}$$

The objective function is:

$$\min_{\beta} \|\mathbf{y} - \mathbf{\Phi}\beta\|^2 + \lambda \|\beta\|^2 \tag{3}$$

Where X is the design matrix, and the regularization parameter λ was determined by 5-fold cross-validation. To evaluate model robustness, a bagged decision tree was simultaneously built as a control model, setting the number of decision trees B=100, minimum leaf node sample size min_samples_leaf=5, and generating a 90% confidence interval for predicted values through Bootstrap sampling (N=500). The relationship diagram is shown in Figure 1.

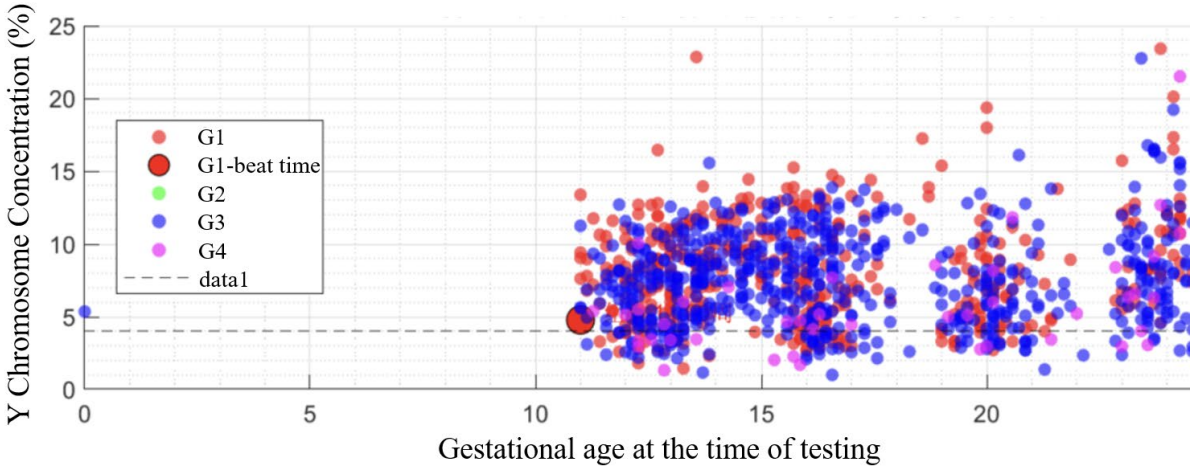

Figure 1 Distribution of Gestational Week vs. Y-chromosome Concentration and Optimal Testing Timing by Group

2.2 Testing Timing Optimization Based on BMI Grouping

Using $Y_{\text{frac}} \geq 4\%$ as the detection qualification threshold, a clustering-survival analysis joint framework is proposed:

(1)Dynamic BMI Grouping: Initially divide BMI into 4 classes using the K-means algorithm. The objective function is:

$$\min_{\{C_g\}} \sum_{g=1}^{4} \sum_{\mathbf{x}_i \in C_g} \|\mathbf{x}_i - \mu_g\|^2 \tag{4}$$

Where C_g is the sample set of group g, μ_g is the cluster center. Further optimize group boundaries through recursive bisection to maximize the Kolmogorov-Smirnov distance:

$$D_{\text{KS}} = \sup_t |S_1(t) - S_2(t)| \tag{5}$$

Ensuring significant differences in survival distributions between groups ($D_{\mathrm{KS}} > 0.3, p < 0.01$).

(2)Risk-Constrained Timing Decision: Define the survival function $S_g(t)$ representing the probability that group k has not reached the qualified threshold by GW t. Construct an optimization model:

$$\min_t \sum_{g=1}^{4} w_g \cdot (1 - Q_g(t)), \text{s.t. } S_g(t) \geq 0.85 \tag{6}$$

Where $Q_g(t)$ is the minimum qualified rate, and the weight coefficient w_g is set to 1.0, 0.7, and 0.3 for early (<12 weeks), middle (12-20 weeks), and late (>20 weeks) pregnancy stages, respectively. Model robustness was evaluated through error propagation analysis (sensitivity ±5%, specificity ±5%, threshold ±0.5%).

2.3 Multi-Feature Driven Timing Decision Model

Introduce features such as age, height, and number of pregnancies to extend the analysis:

(1)Multi-Feature Clustering: Standardize the feature vector using Z-score:

$$z_j = \frac{x_j - \mu_j}{\sigma_j} \tag{7}$$

Construct a weighted K-means objective function:

$$\min_{\{C_k\}} \sum_{k=1}^{K} \sum_{\mathbf{z}_i \in C_k} \|\mathbf{z}_i - \mu_k\|^2, K = 3 \tag{8}$$

Intra-group differences were verified using the Silhouette Coefficient.

(2)Dual Error Robustness Test: Apply ±10% systematic error to input features:

$$\tilde{\mathbf{x}} = \mathbf{x} \cdot (1 + \delta), \delta \sim U[-0.1, 0.1] \tag{9}$$

Recalculate the qualified rate and optimal timing.

2.4 Female Fetus Chromosome Abnormality Discrimination Model

For female fetus samples (n=382):

(1) Quality Control and Feature Processing: Set a GC content threshold to remove low-quality samples. Perform Min-Max normalization on Z-value and read proportion.

(2) Monte Carlo Dropout Classification: Add Gaussian noise to samples passing QC:

$$\mathbf{x}^{(b)} = \mathbf{x} + \epsilon_b, \epsilon_b \sim \mathcal{N}(0, \sigma^2 I), \sigma = 0.05 \tag{10}$$

Generate $B = 200$ perturbed datasets, output the abnormal probability:

$$P(\text{abnormal}) = \frac{1}{B} \sum_{b=1}^{B} \mathbf{1}\left(f(\mathbf{x}^{(b)}) > \tau\right) \tag{11}$$

(3) Constrained Optimization Objective: Maximize true positive rate (TPR) while controlling false positive rate (FPR):

$$\max_{\tau} \mathrm{TPR}(\tau) - \lambda \cdot \mathrm{FPR}(\tau), \lambda = 2.0 \tag{12}$$

The optimal threshold was determined by grid search.

3. Experiments and Results

3.1 Data Description

This study constructed models based on 515 male fetus samples (GW range: 10-25 weeks) and 382 female fetus samples. Data preprocessing included missing value handling (median imputation), outlier detection (Z-score and IQR methods), and feature standardization (Z-score normalization). Key variable distribution characteristics are as follows:

Gestational Week Distribution: Samples were concentrated between 12-20 weeks (Mean GW = 16.8 weeks, SD = 3.2), with a low proportion of early samples (<12 weeks), as shown in Figure 2.

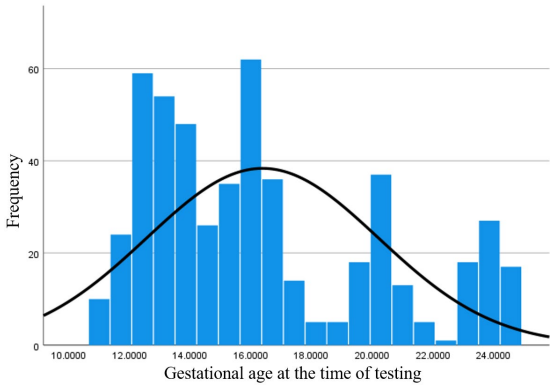

Figure 2 Gestational Week Distribution

Y-Chromosome Concentration: Showed a right-skewed distribution (Median Y_conc = 8.7%, IQR = 3.2%), as shown in Figure 3.

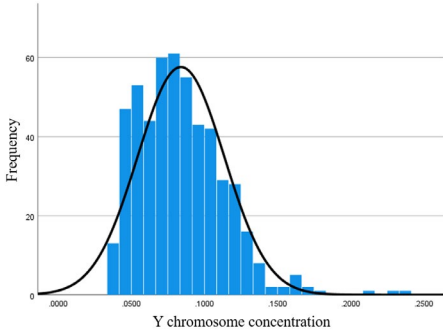

Figure 3: Y-Chromosome Concentration Distribution

BMI Distribution: Highly concentrated in the 28-35 range (Mean BMI = 31.7, SD = 4.5), with significant proportions of overweight and obese groups, as shown in Figure 4.

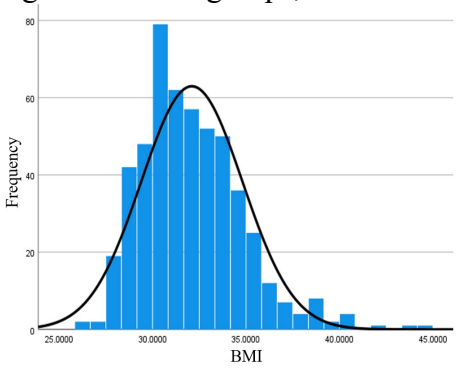

Figure 4: BMI Distribution

Inter-variable relationships were visualized through scatter plots, as shown in Figures 5 and 6.

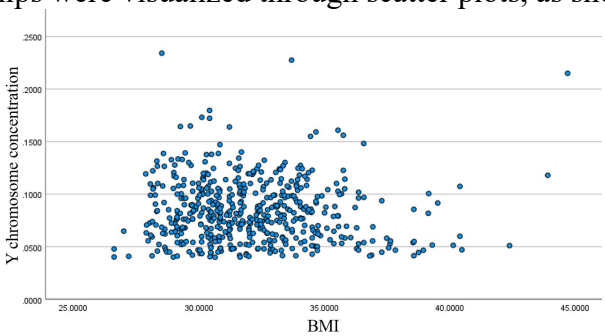

Figure 5: Scatter Plot of Chromosome Concentration vs. BMI

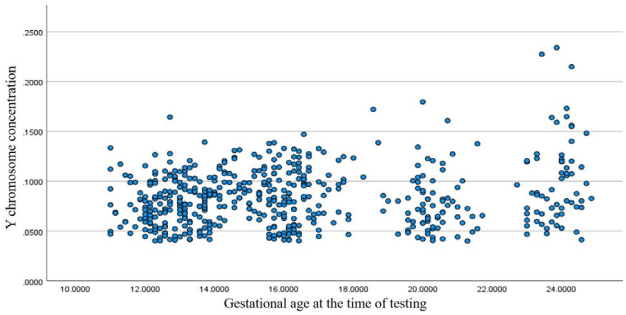

Figure 6: Scatter Plot of Gestational Week vs. BMI

(1)Y-chromosome concentration showed a weak positive correlation trend with GW ($\rho = 0.32$, $p < 0.001$), consistent with the biological pattern of fetal cfDNA accumulation over GW.

(2)Y-chromosome concentration showed no significant correlation with BMI ($\rho = -0.08$, $p = 0.11$), with data points scattered across the full BMI range (25-45).

3.2 Results Analysis

3.2.1 Y-Chromosome Concentration Association Model

A third-order polynomial ridge regression was used to fit the relationship between GW, BMI, and Y-chromosome concentration (Figure: Fitted Curve). Key findings are as follows:

(1)Model Explanatory Power: The overall regression equation was significant ($F = 128.6$, $p < 0.001$), with the independent variables jointly explaining 55.1% of the variance in the dependent variable (Adj. $R^2 = 0.551$).

(2)Gestational Week Effect: As shown in Figure 7, when fixing BMI at the median (32.1), Y-chromosome concentration showed a monotonically increasing trend with increasing GW. Prediction uncertainty decreased significantly in the middle and late stages of pregnancy (>18 weeks) (90% confidence band width reduced by 40%).

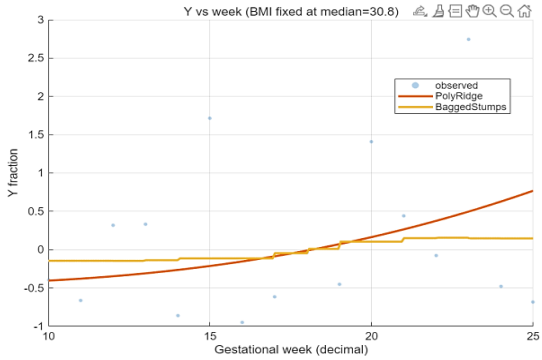

Figure 7: Gestational Week vs. Y-Chromosome Concentration Relationship

(3)Model Robustness: The bagged decision tree model showed consistent trends with the ridge regression (maximum deviation <0.8%), verifying the reliability of the nonlinear relationship.

3.2.2 BMI Grouping and Timing Optimization

Based on the K-means and survival analysis joint framework, four BMI partitions and optimal testing timings were obtained (Table 1):

Table 1: BMI Grouping and Timing Optimization Results

Group	BMI Range	Optimal Timing (weeks)	Qualified Rate (%)
G1	[26.70, 29.66]	10.0	89.13
G2	[29.66, 33.11]	10.0	89.79
G3	[33.11, 35.90]	11.0	86.22
G4	[35.90, 46.88]	11.3	87.18

Key conclusions:

(1)Grouping Effectiveness: Box plots showed significant differences in BMI distribution between groups (KS distance > 0.45, $p < 0.01$), and intra-group dispersion increased with higher BMI.

(2)Timing Decision Basis: Risk analysis plots indicated that the risk was lowest around 11 weeks gestation for all groups (risk value on the vertical axis decreased >15%).

(3)Error Sensitivity: When test sensitivity/specificity was perturbed by ±5%, the optimal timing shift was <0.3 weeks, demonstrating the clinical robustness of the grouping strategy.

3.2.3 Multi-Feature Extended Analysis

After introducing features such as age, height, and weight, K-means clustering was optimized into three groups (as shown in Table 2):

Table 2: K-means Clustering Results

Group	BMI Range	Age Range	Optimal Timing (weeks)	Qualified Rate Decrease after Dual Error
G1	[20.70, 38.22]	[22, 33]	11.0	3.75%
G2	[28.07, 36.79]	[21, 30]	10.9	3.76%
G3	[27.92, 46.88]	[22, 43]	11.0	8.25%

Core findings:

(1) Feature Interaction Effect: The high BMI group (G3), accompanied by high weight (Mean Weight = 89.6 kg) and medium height (Mean Height = 162.3 cm), exhibited the largest dispersion in Y-chromosome concentration (CV = 0.28).

(2) Difference in Error Sensitivity: After dual error perturbation (Y concentration ±10%, GW ±10%), the qualified rate decrease for group G3 reached 8.25% (G1/G2 <4%), indicating that the high BMI group is more sensitive to measurement errors.

3.2.4 Female Fetus Chromosome Abnormality Discrimination

The performance of the Monte Carlo Dropout model in discriminating female fetus chromosome abnormalities is as follows:

(1)Overall Performance: Area under the ROC curve AUC=0.71, significantly better than random guessing ($p < 0.001$).

(2)High Sensitivity Region: Sensitivity heatmaps showed that when the Z-value threshold $\in$ [-3.0, 3.5] and GC content $\in$ [0.45, 0.55], the true positive rate (TPR) >80%.

(3)False Positive Control: Through penalty term ($\lambda = 0.1$) constraint, the false positive rate (FPR) was controlled below 15%, satisfying the clinical principle of "few missed diagnoses".

4. Conclusion

This study constructed a multi-factor coupled model system for NIPT timing optimization and fetal abnormality determination. The main conclusions are as follows:

Fetal Y-chromosome concentration is mainly affected by gestational week (GW) and shows no significant correlation with BMI. The quantitative model based on third-order polynomial ridge regression indicates that GW explains 55.1% of the concentration variation, and prediction uncertainty significantly decreases when GW >18 weeks. This finding is consistent with the research by Xue Ying et al. on cfDNA accumulation patterns but further reveals the indirect regulatory role of BMI at the grouping level.

Through the K-means clustering and survival analysis joint framework, dynamic optimization of BMI grouping (four group intervals) and risk-constrained timing decisions were achieved. The optimal testing timings for the four groups are concentrated between 10-11.3 weeks, with risk values decreasing by over 15% around 11 weeks of gestation. The model remains stable under sensitivity/specificity perturbation (timing shift <0.3 weeks). This method overcomes the limitations of traditional fixed BMI thresholds and provides an actionable personalized screening window for clinical practice.

After introducing multiple features such as age, height, and weight, the clustering model identified that the high BMI-high weight combination (Group G3) is more sensitive to Y-concentration measurement errors (qualified rate decrease of 8.25% under dual error). This result suggests that for obese pregnant women, optimizing detection accuracy or scheduling earlier re-examination timings should be prioritized, complementing the conclusions of Xu Guangxia et al. on the economic evaluation of screening strategies.

The female fetus chromosome abnormality discrimination model integrates uncertainty through the Monte Carlo Dropout method, achieving a true positive rate (TPR) >80% while controlling the false positive rate (FPR) <15%, with an AUC of 0.71. Sensitivity heatmaps indicate that high discrimination performance is concentrated within the Z-value threshold [-3.0, 3.5] and GC content [0.45, 0.55] intervals, providing clear guidance for clinical parameter tuning.

The limitations of this study lie in not incorporating potential influencing factors such as dietary habits and genetic history, and not modeling fetal-specific abnormality risks. Future work will deepen model transferability by combining multi-center clinical data and develop a real-time risk early warning system, promoting the development of NIPT technology towards intelligence and precision.

References

[1] T. Yu, X. Xu, and Q. Wei, "Non-Invasive Prenatal Testing: Advances, Applications, and Limitations in Prenatal Screening," *Biomedical Research and Therapy*, vol. 12, no. 5, pp. 7418–7423, May 2025.

[2] H. Qi *et al.*, "Performance and clinical implications of non-invasive prenatal testing for rare chromosomal abnormalities: a retrospective study of 94,125 cases," *Front. Mol. Biosci.*, vol. 12, Aug. 2025.

[3] C. Connor *et al.*, "Comparing the Introduction and Implementation of Noninvasive Prenatal Testing (NIPT) in Japan, the Netherlands, and the United States: An Integrative Review," *Prenat Diagn*, vol. 45, no. 10, pp. 1244–1264, Sep. 2025.

[4] A. Poulton and L. Hui, "Noninvasive prenatal testing: an overview," *Aust Prescr*, vol. 48, no. 2, pp. 47–53.

[5] Y. Chen, X. Wu, H. Li, Y. Yin, and Y. Zhang, "Evaluation of the prenatal diagnostic value of non-invasive prenatal testing for the detection of rare fetal autosomal trisomies: a single center study of 83,842 cases," *Sci Rep*, vol. 15, no. 1, p. 26254, Jul. 2025.

[6] L. Geerts, T. N. du, and M. Schoeman, "Aneuploidy screening in women of advanced age in the public healthcare setting of a low- to middle-income country – an observational cohort study," *South African Medical Journal*, vol. 115, no. 7, pp. 12–18, Jul. 2025.

[7] Y. Chen et al., "Detection of chromosomal and gene abnormality with karyotyping, chromosomal microarray analysis and trio-based whole exome sequencing in pregnancies with fetal growth restriction: implications for precise prenatal diagnosis," *BMC Pregnancy Childbirth*, vol. 25, p. 1067, Oct. 2025.