Research on Decision-Making for NIPT Testing Timing Based on Mixed-Effects Models and Risk Optimization

Haoyuan Zhang [2,*,#], Jingxiang Qiu [1,#], Luxiang Xu [2], Aoyu Deng [2]

[1] School of Information Science and Engineering, Yunnan University, Kunming, China, 650500

[2] School of Mathematics and Statistics, Yunnan University, Kunming, China, 650500

* Corresponding Author Email: Harryzhanghaoyuan123@outlook.com

[#]These authors contributed equally.

Abstract. Noninvasive prenatal testing (NIPT), which analyzes fetal cell-free DNA in maternal blood to detect chromosomal aneuploidies early in pregnancy, faces challenges in balancing accuracy and timeliness due to the influence of individual variations such as gestational age and body mass index (BMI). To address this issue, this study proposes a dynamic modeling framework integrating clustering and machine learning to optimize the timing of NIPT and improve individualized adaptability. Using fetal Y-chromosome concentration as a core variable, we constructed individual growth curves and mixed-effects models to characterize interactions between gestational age and BMI, combined with Gaussian mixture modeling (GMM) for population stratification. By incorporating survival analysis and risk optimization functions, optimal testing windows were identified for different maternal subgroups, with model robustness verified via Monte Carlo simulations. Results demonstrate that the optimal testing window concentrates between 12–14 weeks of gestation, with the model exhibiting insensitivity to error perturbations. Further integration of multidimensional features—including maternal age, X chromosome concentration, and reproductive history—enabled feature importance analysis through random forest methods, establishing an interpretable high-dimensional clustering prediction framework. This study provides quantitative evidence to support intelligent decision-making in the scheduling and strategy of NIPT.

Keywords: Non-invasive prenatal testing, Mixed-effects model, Gaussian mixture model, Risk optimization.

1. Introduction

Noninvasive prenatal testing (NIPT) identifies fetal chromosomal aneuploidy in early pregnancy by measuring the proportion of fetal cell-free DNA (cfDNA) in maternal peripheral blood. Compared to traditional invasive testing, NIPT offers high safety and accuracy, establishing itself as a vital clinical screening tool [1-2]. In recent years, advancements in mobile health technologies and large-scale retrospective studies have enabled systematic validation of NIPT implementation strategies and performance evaluation across diverse populations [3-4]. However, the stability of test results heavily depends on the relative concentration of fetal cfDNA, which is significantly influenced by multiple factors including gestational age, maternal body mass index (BMI), and individual metabolic variations [5]. Testing too early may result in insufficient fetal cfDNA concentration to meet sequencing thresholds, leading to false negatives. Conversely, delayed testing risks missing the optimal window for intervention. Thus, obtaining reliable results as early as possible while ensuring detection accuracy has become a critical challenge in clinical practice. Determining the optimal testing window amid multidimensional individual variability and constructing dynamic models that balance accuracy and adaptability have emerged as vital research directions for optimizing NIPT strategies [6].

In recent years, advancements in computational methods and data-driven modeling have demonstrated the powerful potential of clustering analysis, mixed-effects models, and risk optimization algorithms in modeling complex medical data. Mixed-effects models (MEM) can simultaneously capture population-level trends and individual variability, finding widespread application in pharmacokinetics, individual growth modeling, and behavioral data analysis.

Clustering algorithms like K-means and Gaussian Mixture Models (GMM) are frequently employed to uncover latent group structures within samples, enabling stratification of individuals across multidimensional features [7]. Additionally, survival analysis methods (Kaplan-Meier and Cox proportional hazards models) have been validated for event-timing modeling and risk-time prediction, reflecting the combined effects of various factors on event occurrence timing. Concurrently, advances in feature importance assessment for machine learning models—such as random forests and SHAP value analysis—offer new approaches to model interpretability and variable selection, aligning algorithmic outcomes with clinical and biological relevance [8]. Existing research demonstrates that integrating clustering, regression, and risk modeling can effectively extract key patterns within multivariate, nonlinear systems, thereby enabling dynamic decision optimization.

Building upon these theoretical and methodological advancements, this paper proposes a dynamic modeling framework based on clustering and machine learning to optimize NIPT detection timing. The study first employs a mixed-effects model to characterize the interaction between gestational age and BMI, using fetal Y-chromosome concentration as the core variable, while incorporating a Gaussian mixture model to identify structural differences among pregnant women. Building upon this foundation, survival analysis and a comprehensive risk function are introduced to construct a risk-minimization model for determining the optimal detection timing. Monte Carlo simulations are employed to assess the model's robustness under detection errors. Furthermore, a random forest approach analyzes the importance of multidimensional features, forming an interpretable high-dimensional clustering prediction framework. By integrating statistical learning with population dynamics modeling, this study explores a feasible pathway for achieving intelligent optimization of detection strategies in biological testing scenarios characterized by significant individual variability.

2. Methods

This paper defines the optimal detection timing problem as a clustering optimization problem. Its objective is to find a mapping function that assigns individuals with multidimensional features (such as BMI, age, etc.) to different subgroups, and to solve for an optimal detection time t* for each subgroup, thereby minimizing the comprehensive risk R(t) composed of the combined risks of non-achievement, delay, and intra-group variation.

2.1. Data Preprocessing

The analysis was based on longitudinal multi-index data obtained from a repository at https://www.mcm.edu.cn/. The dataset comprised information from over 1,200 pregnant women carrying male fetuses. The raw data underwent systematic cleaning and standardization procedures, which included the removal of outliers and appropriate handling of missing values. Pregnancy cycles were uniformly converted into days, and key features such as GC content and gestational age were subjected to range filtering to ensure the overall data quality met the requirements for subsequent modeling.

2.2. Feature Selection

To identify key factors influencing the time required for fetal Y chromosome concentration to reach the target level, this study employed a random forest model for feature importance analysis. Based on ensemble learning from multiple decision trees, random forest effectively evaluates each feature's contribution to the target variable (achievement of target gestational week). Feature importance was calculated using the average reduction in Gini impurity:

$$Importance_j = \frac{1}{N}\sum_T \sum_{t\in T} \Delta\, Gini(t, j) \tag{1}$$

Here, N denotes the total number of trees, T represents the set of all trees, t is a node, and ΔGini (t, j) indicates the reduction in Gini impurity when splitting node t using feature j.

2.3. Cluster Analysis and Group Segmentation

To group pregnant women appropriately, we employed a Gaussian Mixture Model (GMM) to cluster the aforementioned four characteristics. The GMM assumes that the data is generated from a mixture of Gaussian distributions, with its probability density function defined as:

$$p(\alpha) = \sum_{k=1}^{K} \pi_k \cdot \mathcal{N}(\alpha|\mu_k, \Sigma_k) \tag{2}$$

$$\mathcal{N}(x \mid \mu_k, \Sigma_k) = \frac{1}{\sqrt{(2\pi)^d |\Sigma_k|}} \exp\left(-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k)\right) \tag{3}$$

Here, π_k is the mixing coefficient, μ_k and Σ_k represent the mean and covariance of the Gaussian distribution, respectively. Model parameters are estimated using the Expectation-Maximization (EM) algorithm. To mitigate sensitivity to initial values, we employ a hybrid approach combining hierarchical clustering with K-means for initialization optimization. Additionally, a minimum cluster size threshold is set to exclude clusters with insufficient sample size, thereby enhancing clustering stability and representativeness [9].

2.4. Risk Model Construction

2.4.1 Definition of the Comprehensive Risk Function

To evaluate the detection risks across different clustering groups, this paper constructs a comprehensive risk function that optimally balances detection accuracy, clinical timeliness, and population heterogeneity. The function is developed based on the following considerations: NIPT timing decisions inherently involve a multi-objective trade-off, requiring simultaneous mitigation of false negative risk (failure-to-meet-threshold risk) and intervention delay risk (delay risk). Furthermore, to enhance population adaptability, variation risk is incorporated to measure how intra-group individual differences constrain the effectiveness of a uniform testing time point. Therefore, this study integrates the three risks through linear weighting, with the function expressed as follows:

$$\boldsymbol{R(t) = w_1 \cdot P_1(t) + w_2 \cdot P_2(t) + w_3 \cdot P_3(t)} \tag{4}$$

Here, t represents the detection time point (gestational week); w1, w2, and w3 denote the respective weights for each risk factor, satisfying w1 + w2 + w3 = 1. Based on clinical priority, w2 > w1 and w3 are set to reflect a higher penalty for delayed risks.

2.4.2 Definition of the Substandard Risk Function

Definition: The probability that the fetal Y chromosome concentration has not yet reached a reliable threshold (≥4%) when performing NIPT at time point t. This risk reflects the risk of false negatives due to premature testing.

$$\boldsymbol{P_1(t) = \frac{1}{n}\sum_{i=1}^{n} I\,(T_{1,i} > t)} \tag{5}$$

Here,n denotes the total number of pregnant women in the group; T_(at-threshold) represents the gestational week at which the Y chromosome concentration of the i^(th) pregnant woman first reached 4%; I(x) is an indicator function that takes the value 1 when the condition holds and 0 otherwise.

2.4.3 Definition of the Delay Risk Function

Definition: The clinical risk associated with the delay in detection at time point t relative to an ideal baseline time point. Delayed detection may result in missed intervention windows, increased treatment difficulty, and other consequences.

$$\boldsymbol{P_2(t) = \frac{\max(0, t - t_0)}{T_{\max} - t_0}} \tag{6}$$

Here, t0 is the ideal earliest detection time, set to 10 weeks; Tmax is the latest permissible detection time, set to 25 weeks; the denominator is used for normalization, ensuring the delay risk falls within the [0, 1] interval.

2.4.4 Definition of the Variation Risk Function

Definition: Reflects the degree of variation in the time to reach target level T among pregnant women within a group. Greater variation indicates poorer uniformity at a common testing point and a higher necessity for individualized recommendations.

$$\boldsymbol{P_3 = \frac{\sigma_{T_1}}{\mu_{T_1}}} \tag{7}$$

Here,$\sigma_{T_1}$is the standard deviation of the time to meet standards among pregnant women in the group; is the mean time to meet standards among pregnant women in the group; $\mu_{T_1}$is the coefficient of variation, used to measure the relative dispersion.

2.5. Solving for the Optimal Detection Time

The optimal detection time t* is determined by minimizing the comprehensive risk function within the feasible time interval:

$$\mathbf{t^* = arg \min_{t\in[10,25]} R(t)} \tag{8}$$

To assess the model's robustness against detection errors, we employed a Monte Carlo method to simulate the impact of measurement errors. The core idea is to repeatedly introduce random perturbations—assuming errors follow a normal distribution—into the original feature data, recalculate the optimal solution, and then quantify the variability in outputs such as the optimal timing t*. This study validated the model's stability by calculating the standard deviation of t* through 1000 repeated samples.

3. Results and Discussion

3.1. Feature Importance Analysis Results

We performed feature importance analysis on all variables using the random forest method (Figure 1), identifying the following as relatively important features: BMI (0.42), X chromosome concentration (0.32), age (0.20), and number of deliveries (0.06).

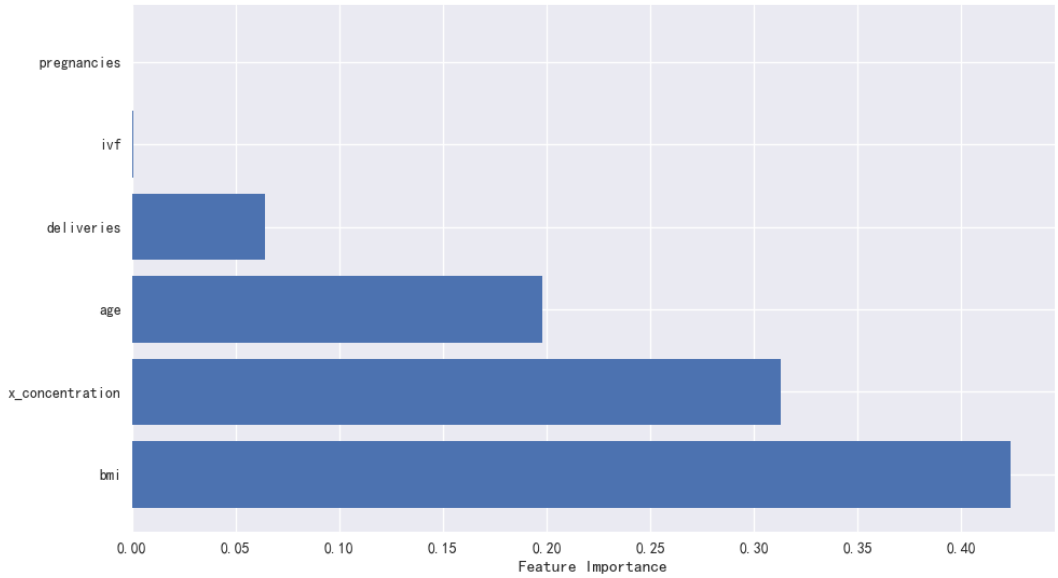

Figure 1. Random Forest Feature Importance

3.2. Analysis of Multidimensional Clustering Results

Based on a Gaussian mixture model (GMM), four key characteristics—BMI, X chromosome concentration, age, and number of births—were clustered. The composite scoring model determined the optimal number of clusters to be 6. The range of characteristics and sample size distribution for each cluster are shown in Table 1.

Table 1. Multidimensional Clustering Results of Pregnant Women Based on Gaussian Mixture Models and Optimal Screening Timeline

Group	BMI Range	X Chromosome Concentration	Age Range	Number of Pregnancies	Group Size	Optimal Timing
1	28.1,35.5	0.1,0.2	21.0,30.0	0.0,0.0	31	13.29 weeks
2	28.9,35.7	0.0,0.1	23.0,4.01	2.0,3.0	16	13.19 weeks
3	27.9,31.9	0.0,0.1	30.0,41.0	1.0,1.0	14	13.14 weeks
4	26.6,46.9	−0.1,0.1	21.0,38.0	0.0,0.0	144	13.71 weeks
5	27.2,33.4	−0.0,0.1	24.0,32.0	1.0,1.0	31	11.71 weeks
6	34.7,41.1	−0.0,0.1	22.0,36.0	1.0,1.0	14	12.29 weeks

The clustering results indicate that GMM successfully divided the pregnant women population into subgroups with distinct physiological profiles. For instance, Cluster 1 was characterized by moderate BMI, higher X chromosome concentration, and young primiparous women; whereas Cluster 3 exhibited lower BMI, lower X chromosome concentration, and advanced maternal age. This outcome validates the effectiveness of GMM in handling multidimensional, non-spherical distribution data, revealing more complex population structures beyond the single BMI metric.

3.3. Determination of Optimal Detection Time and Risk Assessment

By minimizing the composite risk function, the optimal NIPT testing time was determined for each cluster, with results shown in Table 1.

Analysis revealed significant variations in optimal detection timing across clusters, ranging from 11.71 to 13.71 weeks. This divergence directly reflects shifts in the minimum risk function points due to differing feature combinations within each group. For instance, Cluster 5 achieved the earliest optimal detection time (11.71 weeks), indicating that its feature profile enables earlier attainment of reliable detection criteria, thereby effectively shortening the intervention window. In contrast, Group 4 exhibited the latest optimal detection time (13.71 weeks), suggesting this cohort may require extended waiting periods to ensure detection accuracy. This demonstrates the value of multidimensional clustering models in achieving personalized, precise timing recommendations, aligning with observed variations in NIPT detection efficacy across different populations in clinical studies [10].

3.4. Robustness Analysis of the Model

The results indicate that within the error range, the average fluctuation range of the optimal detection time falls between [−0.045, 0.082] weeks, while the average variation range of the comprehensive risk lies between [-0.37, 0.08] (Figure 2). Both fluctuations exhibit minimal amplitude, falling within acceptable engineering application ranges. This outcome demonstrates that the risk optimization model constructed in this study exhibits robust stability and insensitivity to input data noise. This provides reliable assurance for its application in real clinical settings, aligning with recent research conclusions regarding NIPT testing stability [7].

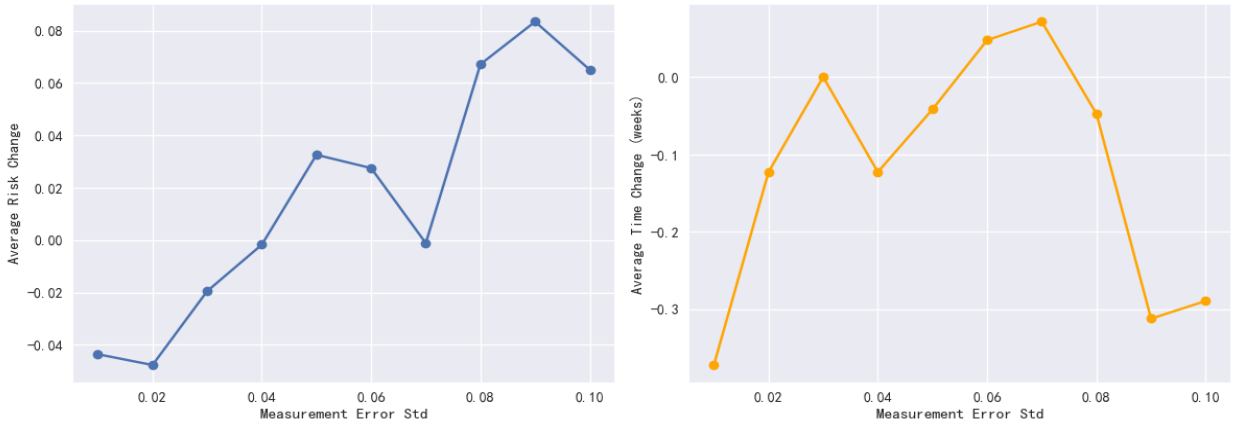

Figure 2. Impact of Measurement Error on Various Indicators

4. Conclusions

This study developed a computational framework integrating data-driven approaches and model optimization to determine the optimal testing window for male fetuses in non-invasive prenatal testing (NIPT). The framework systematically combines core algorithms—including random forest feature selection, Gaussian mixture model multidimensional clustering, and comprehensive risk function optimization—to achieve refined stratification of the pregnant population. Results demonstrate that multidimensional features—including BMI, X chromosome concentration, maternal age, and parity—effectively partition pregnant women into six distinct subgroups. This approach overcomes limitations of traditional single-factor grouping (e.g., BMI) and reveals more complex intrinsic population structures. By minimizing the comprehensive risk function for each subgroup, the model outputs differentiated optimal detection timings (ranging from 11.71 to 13.71 weeks of gestation), fully validating the necessity and feasibility of personalized recommendation strategies based on refined grouping. Furthermore, Monte Carlo simulation analysis demonstrated the model's robust resilience to random errors in detection data. Both the optimal testing time and the fluctuation range of the comprehensive risk remained within clinically acceptable limits, ensuring the model's application potential and stability in real-world clinical settings.

In summary, the proposed method provides a reliable technical pathway for precision optimization of NIPT testing timing. It not only enhances the scientific rigor and specificity of testing strategy formulation but also offers crucial algorithmic support for transitioning clinical practice from a “one-size-fits-all” to an “individualized” management model. Future research directions include external validation across broader populations and exploring the integration of dynamic monitoring data to further enhance the model's predictive performance. Additionally, combining systematic evaluation with implementation science research will optimize strategies for integrating NIPT into diverse clinical pathways, ultimately delivering actionable decision support tools for personalized NIPT testing.

References

[1] Shi Weihui, Xu Chenming. Application Value of Non-Invasive Prenatal Testing in Diagnosing Maternal Complications and Concomitant Conditions in Obstetrics [J]. Journal of Practical Obstetrics and Gynecology, 2025, 41(08): 617-620.

[2] Sebire E, Rodrigo H C, Bhattacharya S, et al. Correction: The implementation and impact of non-invasive prenatal testing (NIPT) for Down's syndrome into antenatal screening programmes: A systematic review and meta-analysis. [J]. PloS one,2025,20(2): e0318985.

[3] Bendsen B S, Skinner C T, O’Reilly L S, et al. Exploring engagement patterns within a mobile health intervention for women at risk of gestational diabetes. [J]. Women’s health (London, England), 2025, 2117455057251327510.

[4] Shen S, Qi H, Yuan X, et al. The performance evaluation of NIPT for fetal chromosome microdeletion/ microduplication detection: a retrospective analysis of 68,588 Chinese cases. [J]. Frontiers in genetics, 2024, 151390539-1390539.

[5] Yamamoto K, Suzumori N, Miura K, et al. Clinical Implications of Low Cell-Free DNA Fetal Fraction in Non-Invasive Prenatal Testing: A Retrospective Cohort Study of 40,716 Pregnancies. [J]. Prenatal diagnosis,2025

[6] Yu Dandan, Li Fengjin, Yao Xinyu, et al. Efficacy Analysis of NIPT in Detecting Common Fetal Chromosomal Aneuploidy in Different Precautionary Groups [J]. Chinese Journal of Family Planning and Obstetrics & Gynecology, 2025, 17(03): 72-76+82. DOI: CNKI:SUN:JHFC.0.2025-03-016.

[7] Wang Y, Liu M, Gao Z, et al. Detection of genomic variants by genome sequencing in foetuses with central nervous system abnormalities. [J]. Annals of medicine,2024,56(1):2399317.

[8] Judi A A, Suresh P, Raj B A T. FB-UNet++: federated biometric UNet++ model for segmentation and classification network of fetal anomaly detection in prenatal care[J]. Biomedical Signal Processing and Control,2026,112(PC):108706-108706.

[9] Luong A L T, Nguyen A D, Dao T T, et al. Combined Cell-Free DNA Screening for Aneuploidies and Selected Single-Gene Disorders for Pregnancies with Sonographically Detected Fetal Anomalies: Detection Rate and Residual Risk. [J]. Prenatal diagnosis,2024,45(1):70-76.

[10] Wang C, Zhang H, Liu Y, et al. Results and indications changes by prenatal diagnosis of 3,458 pregnant women: a 10-year retrospective study. [J]. Annals of medicine,2025,57(1):2563753.