BMI-Stratification-Driven Optimization of NIPT Timing and Fetal Abnormality Determination

Ruijing Mei *, Zhiyuan Yao, Qiong Cai

School of Artificial Intelligence, Jianghan University, Wuhan, China, 430056

* Corresponding Author Email: mrj678886@gmail.com

Abstract. Noninvasive prenatal testing (NIPT) analyzes cell-free fetal DNA fragments in maternal blood and faces two key challenges: precisely selecting the testing time point and efficiently determining fetal chromosomal abnormalities. Integrating generalized linear modeling and clustering, this study focuses on BMI-driven optimization of NIPT timing and the determination of fetal anomalies. First, we employ a generalized linear regression model with Pearson correlation analysis to quantitatively characterize the associations between fetal Y-chromosome concentration and indicators such as gestational age and BMI. The results indicate that the quadratic BMI term exhibits a pronounced nonlinear inhibitory effect;evaluation based on GLM metrics shows that the model stably captures the positive effect of gestational age and the overall suppressive effect of BMI. Next, this article combines K-means clustering with decision-tree regression to stratify maternal BMI among pregnancies with male fetuses, yielding a grouping scheme with within-cluster homogeneity of 92%, and construct a composite risk function R by integrating window-period risk and accuracy risk to determine the optimal time point. Finally, this article adopts the pipeline “fit a Gaussian mixture model and use the midpoint c of the two component means as a data-driven threshold→ multi-model cross-validation for model selection→ gestational-age grid computation→ use τ to obtain the earliest attainment time T*→ Pareto frontier trade-off, ”and, with K-means–based BMI stratification and τ sensitivity analysis, shows that the recommended time point exhibits notable robustness.

Keywords: Fetal abnormality determination, Generalized linear regression, K-means clustering, Gaussian mixture model.

1. Introduction

This study develops a BMI-stratified approach to personalize the timing of NIPT, enhancing detection accuracy and reducing repeat blood draws by adaptively determining optimal testing schedules for different BMI subgroups. Shi Weihui and Xu Chenming [1] noted that NIPT, based on maternal peripheral blood cfDNA and next-generation sequencing, offers higher sensitivity and specificity for trisomy 21/ 18/ 13 screening and has become the first-line option; the new generation of NIPT has expanded to aneuploidies beyond autosomal trisomies, microdeletions/microduplications, and dominant monogenic disorders, and shows potential for identifying and managing maternal complications such as gestational hypertension, gestational diabetes mellitus(GDM), and fetal growth restriction(FGR). They systematically reviewed its current clinical applications. Jiang Liya et al [2] indicated that NIPT, which evaluates genetic disease risk by sequencing cfDNA of fetal origin present in maternal peripheral blood, confers advantages over traditional methods;its clinical application continues to deepen and it occupies a central role in prenatal screening. Their article reviews technological developments and clinical comparisons. Zhang Peng et al [3] reported that, in a retrospective dataset of 17, 211 cases from 2005-2023, after the implementation of NIPT the detection rate of sex-chromosome aneuploidies increased from 0.35%to 1.35%, with positive associations to maternal age and testing volume; high-risk NIPT results became a principal indication for prenatal diagnosis; follow-up showed relatively high termination rates for certain categories, underscoring the need to strengthen ethical and genetic counseling support. Zhang Zhenzhen et al [4] pointed out that combining NIPT with concurrentβ-hCG and PAPP-A at 9 to 13+ 6 weeks of gestation can significantly improve the detection of chromosomal abnormalities in early pregnancy, with the three-way combination demonstrating better concordance and detection efficiency than any single test. Yang Peng et al [5] found that, among advanced-maternal-age pregnancies, ultrasound combined with

NIPT significantly improves sensitivity, specificity, and positive predictive value (PPV) for chromosomal abnormalities compared with either modality alone, highlighting the benefit of multimodal integration.

Most existing studies remain at the level of technical overviews, combined screening strategies, or overall performance evaluations; they have not established models that adapt the choice of testing time point, repeat-sampling strategies, and abnormality-determination thresholds to BMI variation. Moreover, study designs are often single-center or retrospective, lacking prospective multicenter validation and decision-curve evaluation aligned with clinical pathways, which limits the provision of operational, quantitative evidence for individualized scheduling of NIPT.

In this work, we integrate generalized linear regression with Pearson correlation analysis to model the relationships between gestational age/BMI and Y-chromosome concentration; combine K-means clustering with decision-tree regression to achieve BMI stratification and prediction of the optimal testing time; and employ GMM-based threshold labeling together with Logistic/GBDT/Random Forest models selected via five-fold AUC, plus Pareto frontier analysis and sensitivity analysis, to form a robust recommendation pipeline. Addressing the challenges of individualized NIPT scheduling and anomaly determination, the method yields a BMI-adaptive testing time $T*$, gestational-age–dependent attainment-probability curves, and a threshold-calibration scheme.

Data source:https: //cumcm. cnki. net/cumcm/studentHome/studentHome (Problem C).

2. Study of Generalized Linear Modeling and Pearson Correlation on Factors Affecting Fetal Y-Chromosome Concentration

2.1. Model Construction

Fetal Y-chromosome concentration is influenced by multiple factors, including gestational age, BMI, age, and content measures, among others. Because the standardized concentration data may contain negative values, this study adopts a generalized linear model(GLM)with a log link. This specification effectively accommodates non-normally distributed response variables and uses the link function to relate covariates to the response. The core rationale is as follows:

After standardization, the concentration column contains negative values and thus cannot directly employ a log link. Therefore, we first apply a monotonic shift to the concentration:

$$Y_{pos} = Y + 3.00529 \tag{1}$$

After the shift, all observations$Y_{pos}>0$ satisfy the GLM requirement that the response be strictly positive. Y denotes the fetal Y-chromosome concentration.

Using the Gaussian distribution family and the logarithmic link function, the model is defined as:

$$g(\mu_i)=\log(\mu_i)=\beta_0+\beta_1\cdot t_i+\beta_2\cdot BMI_i+\beta_3\cdot BMI_i^2+\beta_4\cdot GC_i+\beta_5\cdot W_i \tag{2}$$

Among them, $\mu_i=E(Ypos|X_i)$represents the conditional mean of Ypos given the independent variable as Xi, β0 represents the intercept term， β0 is the intercept, and β1 to β5 are the coefficients of the independent variables. t represents age, and W represents gestational age (weeks).

The Pearson correlation coefficient, a statistic measuring linear association between two continuous variables[6], is used to explore correlations between Y-chromosome concentration and indicators such as gestational weeks and BMI. The formula is as follows:

$$r_{XY} = \frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}\cdot\sqrt{\sum_{i=1}^{n}(Y_i-\overline{Y})^2}} \tag{3}$$

2.2. Model Estimation

Parameter estimation: In the GLM framework, parameter estimation aims to identify the parameters that maximize the likelihood of the observed data[7]. Specifically, this includes the

intercept β0 and regression coefficients β1, β2, …By constructing the log-likelihood function and solving for its extremum, parameter estimates are obtained. Implementation is conducted in Python using the statsmodels library, which automatically performs the iterative solution to ensure accuracy.

Significance testing: To determine whether model and variable effects exceed random error, significance testing is performed. For the overall model and individual coefficients, the t-test statistic is applied.

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \sim t(n - p - 1) \tag{4}$$

If the corresponding p-value is less than the significance level p=0. 05, the variable is deemed to have a statistically significant effect on Y-chromosome concentration.

Goodness-of-fit evaluation: Within the GLM framework, McFadden's pseudo R^2, deviance explained, and AIC/BIC are used to assess fit.

McFadden's pseudo R^2:

$$R^2_{McF} = 1 - \frac{\ln L_{model}}{\ln L_0} \tag{5}$$

Deviance explained:

$$\text{Explained Deviance} = 1 - \frac{D_{model}}{D_0} = 1 - \frac{\text{Residual Deviance}}{\text{Null Deviance}} \tag{6}$$

Information criteria:

$$AIC = 2k - 2\ln L_{model}, BIC = k\ln n - 2\ln L_{model} \tag{7}$$

(Where kis the number of parameters and n is the sample size.)

2.3. Data visualization and result analysis

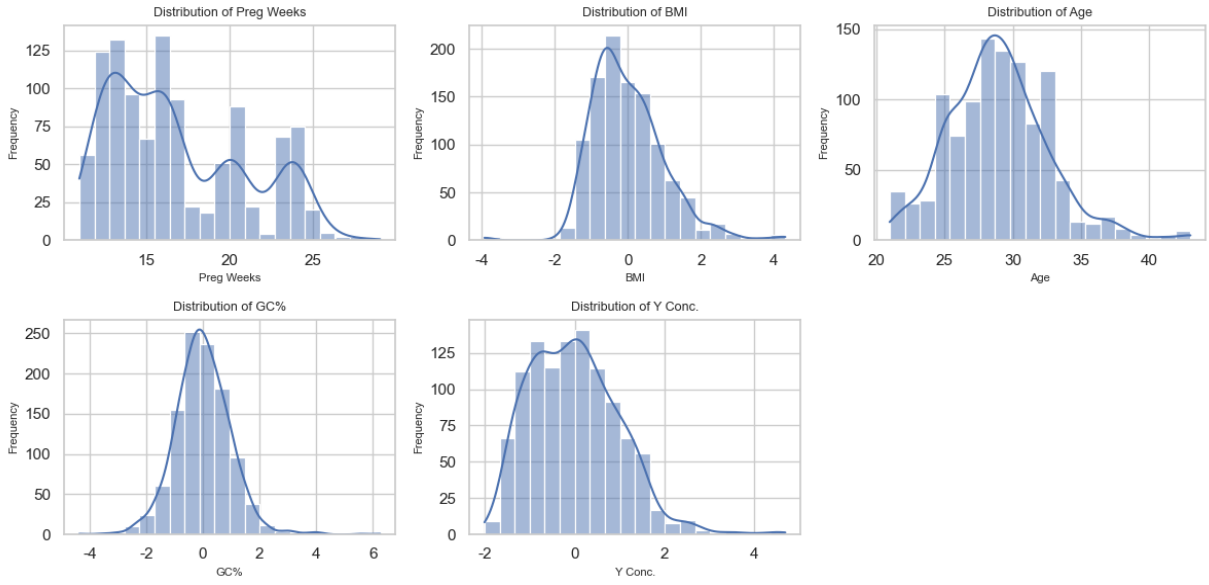

Figure 1. Histogram of variable distribution and kernel density curve

As shown in Figure 1, the samples are concentrated in typical gestational weeks(such as 12-16 weeks, 20-25 weeks), suggesting a structural distribution of "week measurements"; The age is close to normal and concentrated at 28-32 years; The GC content distribution is concentrated with few extreme values, which conforms to the characteristics of sequencing quality control variables.

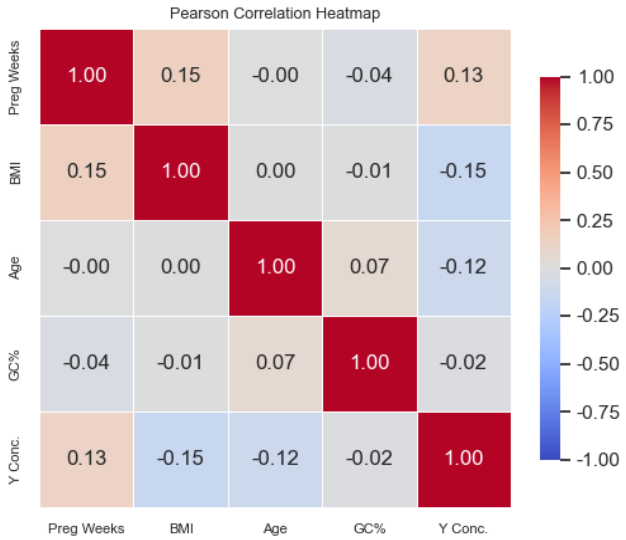

Figure 2. Pearson correlation heat map

As shown in Figure 2, the correlation coefficients between gestational weeks, BMI, age, GC content and Y are+0. 13, -0. 15, -0. 12, -0. 02, respectively; The correlations among the independent variables were generally low, indicating a lower risk of collinearity. Then a two-dimensional relationship review was conducted, and scatter plots of Y with their respective variables were plotted to determine whether it was linear or nonlinear, outliers, and heteroscedasticity characteristics. The results are as follows:

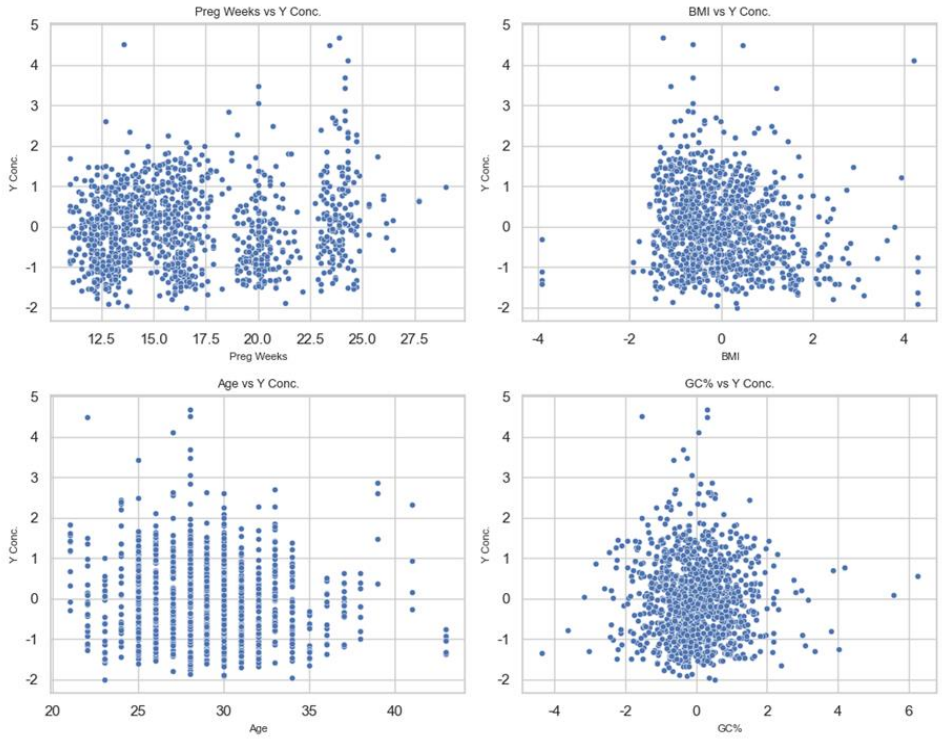

Figure 3. Scatter plot of the independent variable and Y concentration

As can be seen from Figure 3, gestational week-Y chromosome concentration shows a slightly upward trend, but due to the influence of"zonal bands", it is overall closer to segmental linearity rather than strict linearity;Bmi-y chromosome concentration shows a slightly negative relationship, and the distribution of Y chromosome concentration is generally"downward"when BMI is higher; Age-y chromosome concentration clouds are dispersed and there is no definite monotonic tendency; GC content-Y chromosome concentration is basically a circular cloud, and the first-order effect is very weak.

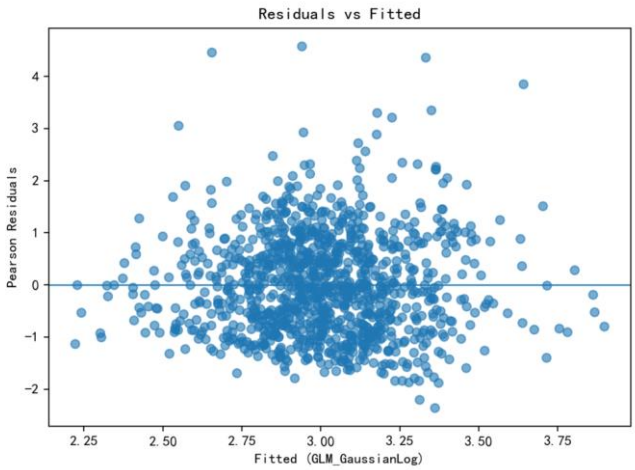

Figure 4. Residuals versus fitted values graph

To verify the fitting rationality of the generalized linear regression model, it can be seen fromFigure4 that the residuals are randomly distributed on both sides of 0, and there is no systematic pattern or severe heteroscedasticity overall. This indicates that the established model is basically in line with the initial assumptions and the fitting is reasonable.

3. K-means and decision tree regression study for the grouping of pregnant women'BMI and the time point optimization of NIPT

The core of this study is to determine the optimal NIPT time points for different groups of pregnant women under the premise of finding reasonable BMI groups for pregnant women, in order to reduce potential risks. Clinically, the time to reach the target concentration of Y chromosome in male fetuses is significantly influenced by BMI;High BMI often reduces the relative proportion of free DNA in the fetus, and reaching the same probability of attainment usually requires a later test time point; In contrast, those with a lower BMI reach the target earlier and can be tested moderately earlier within a reasonable range. Based on this, this paper uses a two-tier model of"cluster-regression": first, group by BMI using k-means(K=4); Then within each group, a decision tree regression was constructed with gestational weeks as the regression target and pregnant women'age, BMI, etc. as independent variables to output the recommended test time points.

3.1. Model Construction

K-means clustering aims to partition samples into K groups so that within-cluster similarity is maximized while between-cluster dissimilarity is maximized[8]. In the present study, the focus is on the relationship between standardized maternal BMI values and Y-chromosome concentration. The core principle is minimization of the sum of squared errors(SSE): by iteratively updating cluster centroids, the algorithm minimizes the sum of Euclidean distances from each sample to the centroid of its assigned cluster.

Data preprocessing:First, the BMI data of mothers carrying male fetuses are collected and standardized to eliminate the effect of measurement units;the formula is:

$$\text{Standardized BMI}=\frac{\text{Raw BMI}-\mu}{\sigma} \tag{8}$$

Where μ denotes the mean BMI and σ the standard deviation.

Selection of the number of clusters:The elbow method is employed to identify the optimal KKK. By computing the sum of squared errors(SSE)for different values of KKK, the elbow method selects the point at which the SSE's rate of decrease markedly slows(the"elbow")[9]. The SSE is defined as:

$$SSE=\sum_{i=1}^{K}\sum_{x\in C_i}\left\|x-\mu_i\right\|^2 \tag{9}$$

where Ci denotes the i-th cluster, μi the centroid of that cluster, and x a sample point(standardized BMI). A lower SSE indicates smaller within-cluster distances and, consequently, a more reasonable partition.

Clustering execution: The K-means algorithm is applied to partition the standardized BMI data into K groups(with K=4 in the present study). The algorithm iterates until convergence.

Groupwise analysis: The clustering partitions pregnant women into four groups, each corresponding to a BMI range(from low to high), and an association analysis is conducted with Y-chromosome concentration.

Decision-tree regression performs recursive partitioning on input features(e. g. , BMI, age, height)to predict a continuous outcome[10]. The principle is based on information gain or reduction in mean squared error: at each split, the optimal feature and threshold are selected so that the variance of the target variable(gestational week)within the resulting subsets is minimized. This approach not only accommodates interaction effects among features(such as the joint influence of BMI and age on the optimal time point)but also provides individualized time-point predictions for each BMI group, thereby reducing the risk of missed detection.

Feature engineering: Input features include BMI, maternal age, and height(continuous variables); the output variable is the optimal NIPT time point(gestational week), defined based on historical data.

Tree construction: Starting from the root node, recursively select the optimal feature and split point; the formulation is as follows:

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{10}$$

Where y_i denotes the actual gestational week and$\hat{y}_i$the predicted gestational week.

Model training: The decision tree is fitted on the training dataset, and hyperparameters are optimized via cross-validation.

Prediction output: The final model yields the following groups—Group 0(BMI range:(−∞, 31. 01]; Group 1(BMI range:(31. 01, 33. 90];Group 2(BMI range:(33. 90, 37. 90](33. 90, 37. 90](33. 90, 37. 90]);Group 3(BMI range:(37. 90, +∞). Among these, Group 0 exhibits the highest mean Y-chromosome concentration(0. 8 AU), whereas Group 3(high BMI)shows the lowest(0. 5 AU). A plausible mechanism is the dilution effect of BMI:individuals with higher BMI tend to have larger blood volumes and a lower fraction of fetal DNA, leading to weaker Y-chromosome signals. Grouping effectiveness is supported by SSE(at K=4, SSE=120, with a slower decline relative to K=3), indicating that four clusters are optimal.

The predicted optimal NIPT time point is explicitly BMI-dependent:Group 0 is recommended for slightly later testing(15. 2 weeks), as strong signals at low BMI imply that very early testing may be resource-inefficient; Group 3 should be tested earlier(10. 8 weeks)to mitigate the risk of false negatives associated with high BMI. Maternal age and height act as secondary features:when age>35>35>35 years, the recommended time point shifts by+1+1+1 week, while height shows a weak effect. The model fit of R2=0. 85R^2=0. 85R2=0. 85 indicates strong explanatory power.

3.2. Data Visualization and Results Analysis

Figure 5 reveals significant differences in Y-chromosome concentration across maternal BMI groups, which may suggest that the influence of BMI on fetal sex differs. Accordingly, pregnant women with higher body weight may require different testing time points.

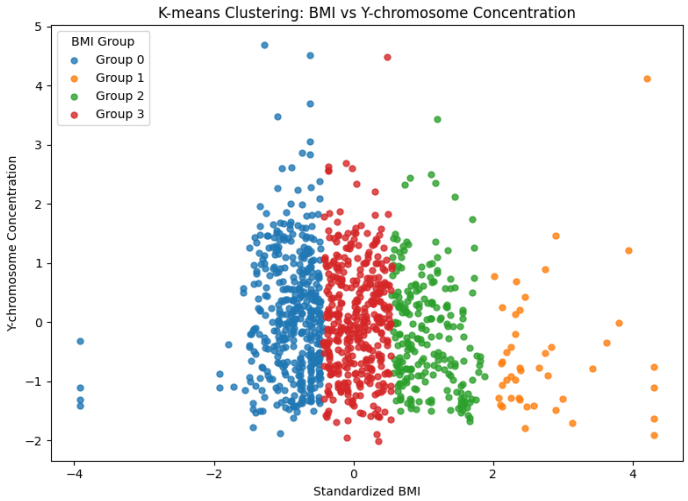

Figure 5. K-means clustering results.

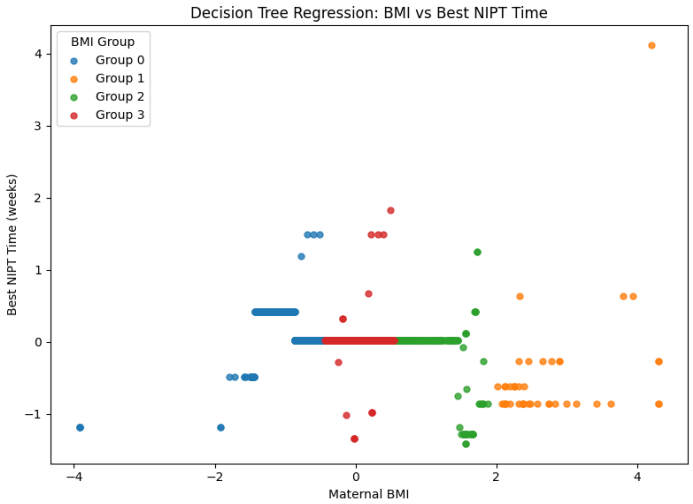

Figure 6. Decision-tree regression results

By further analyzing maternal BMI, age, and height, the optimal NIPT testing time (gestational age) is predicted;the decision-tree regression results shown in Figure 6 further substantiate the aforementioned findings.

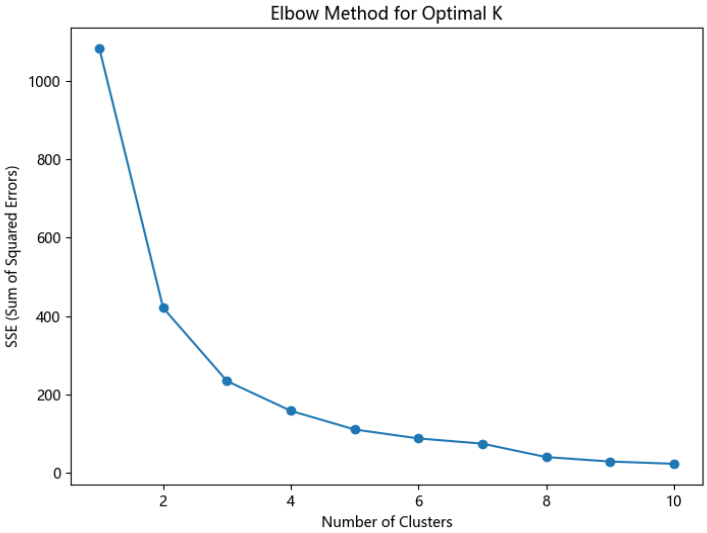

Figure 7. Elbow-method plot

As shown in Figure 7, when the number of clusters K=4, the decline in SSE slows markedly;therefore, selecting four clusters is justified in the present study. From the figure, it can be observed that the reduction in SSE begins to level off at K=4, indicating that four clusters constitute the optimal choice.

4. Study on GMM Thresholding of NIPT Attainment Probability and Ensemble Classification Modeling

4.1. Model Analysis

This chapter adopts a“multi-factor—coordinated optimization”framework:while considering maternal age, height, weight, and other features, optimization of the testing time focuses on the determination of“whether the passing standard is met at a given gestational week. ” Concretely, a Gaussian mixture model(GMM, K=2)is first fitted to the one-dimensional distribution of Y-chromosome concentration to learn a data-driven threshold ccc, on the basis of which observations are labeled“pass/fail”[11]. Subsequently, among Logistic, GBDT, and Random Forest classifiers, the model with the best performance is selected via five-fold cross-validated AUC and used to estimate the“attainment probability given gestational week and individual features. ”To reduce within-group heterogeneity, BMI is further stratified using K-means(K=4). Within each BMI group, the average attainment-probability curve over gestational week is computed along a gestational-age grid, and the earliest recommended time point satisfying preset thresholds(e. g. , 0. 80/0. 85/0. 90)is identified. Finally, a Pareto frontier is used to visualize the trade-off between“earlier testing”and“higher attainment probability, ”facilitating clinically balanced choices.

4.2. Model Building

Threshold and Labeling(GMM)Model the Y concentration on male fetal samples using GMM(K=2K=2K=2), record the mean of the two components asμ1, μ2, and take it as the data-driven threshold. $c = \frac{\mu_1+\mu_2}{2}$generated:pass=1 if $Y \geq c$, otherwise pass=0.

The candidate models predicted by probability are Logistic, GBDT, RF;Model selection based on AUC metrics within a five-fold hierarchical and a cross-validation framework[12]. $\hat{\Pr}(\text{pass} \mid t, X)$Features included gestational age t, BMI(and BMI²), age, etc. Get the best model.

BMI stratification(K-means)K-means clustering is performed on BMI to obtain four identical proton groups, and the clustering boundaries are de-standardized to the original unit(kg/m²)to form intervals of"group 0-3".

The gestational week-pass rate curve within each BMI group g and the recommended time point along the gestational grid t∈[11, 29]weeks, set the"gestational week"of the sample to t, calculate the probability of reaching the standard for each individual using the optimal classifier and take the average to obtain.

$$\bar{p}_g(t) = \frac{1}{|S_g|}\sum_{i\in S_g} \hat{\Pr}(\text{pass} \mid t, X_i) \tag{11}$$

Set the thresholdτ∈{0. 80, 0. 85, 0. 90}for the recommended gestational weeks.

$$T^* = \arg\min\{t: \bar{p}_g(t) \geq \tau\} \tag{12}$$

If the threshold is not reached within the interval, take the peak point as an alternative$\bar{p}_g(t)$.

Pareto Frontier and Sensitivity: Non-dominated screening with target pairs (t, $\bar{p}$g(t), plot the $T^{\backslash *}$ Pareto frontier to show a compromise of " early detection/guaranteed pass rate"; And recalculate and at τ=0.80/0.85/0.90 to assess the robustness $\bar{p}_g(T^{\backslash *})$ of the conclusion to threshold selection.

4.3. Model Solution

Preparation and stratification: After data cleaning, apply K-means clustering to BMI(K=4), and record within-group sample sizes and boundary cut points.

Thresholds and labels:Learn the threshold c via a Gaussian Mixture Model(GMM)and generate“pass/fail”labels.

Model training: Compare Logistic, GBDT, and Random Forest classifiers using five-fold AUC;select the best model and retrain it on the full sample to obtain predicted“attainment probabilities. ”

Curves and time points: Within each BMI group, sweep a gestational-age grid to compute the average attainment-probability curve, and determine the earliest recommended time point at thresholds 0. 80/0. 85/0. 90; when necessary, replace with the peak point.

Pareto frontier and visualization: Plot the Pareto frontier and the“per-group curves+recommended time points”figure to present the trade-off between“earlier”and“more stable”recommendations.

Evaluation and robustness: Report the cross-validated AUC of the selected classifier; provide a sensitivity table of recommended time points across thresholds; optionally construct bootstrap intervals for the curves and time points or conduct noise-perturbation comparisons to demonstrate stability of the conclusions.

4.4. Data Visualization and Results Analysis

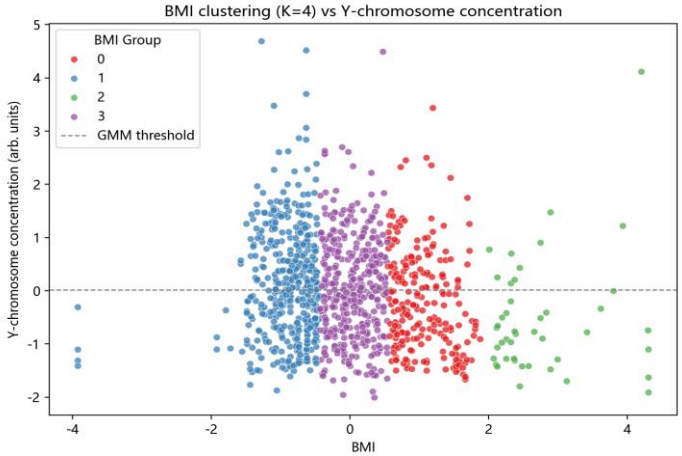

Figure 8. BMI clustering and scatter plot of Y-chromosome concentration

From Figure 8, the high-BMI group(Group 3)exhibits overall lower Y-chromosome concentration, whereas the mid-low BMI groups(Groups 0 and 2)show a more concentrated distribution with fewer outliers. Accordingly, the figure suggests a negative association between BMI and Y-chromosome concentration.

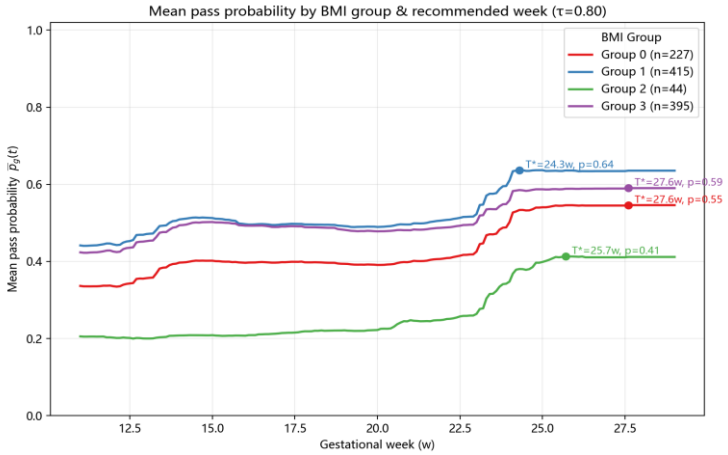

Figure 9. Average attainment-rate curves and recommended gestational weeks by BMI group

The average attainment-rate curves and recommended gestational weeks are plotted for each BMI group. As shown in Figure 9, the curves rise in the later period(26–29 weeks)and then gradually stabilize;the curve for Group 3(highest BMI)is overall lower, with a peak below that of the other groups. Under the settingτ=0. 85, the four groups yield T$*$at approximately 24. 3/27. 6/27. 6/25. 7 weeks, respectively, indicating that higher BMI warrants earlier attention.

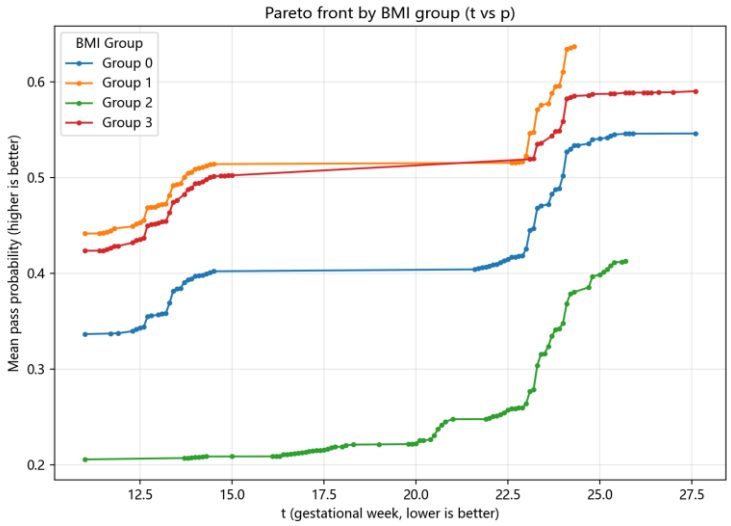

Figure 10. Pareto frontier plot.

From Figure 10, the variations in the attainment probability of Y-chromosome concentration and time efficiency across different stratification schemes can be analyzed, as can the extent to which various factors influence the testing time point.

5. Conclusion

This study addresses the clinical challenges of individualized scheduling for NIPT and anomaly determination, and proposes an integrated approach of “correlation characterization—stratified modeling—probability recommendation—robust validation.” First, using generalized linear regression(log link)combined with Pearson correlation analysis, it quantitatively elucidates the relationships among gestational age, BMI, and fetal male Y-chromosome concentration;second, it performs data-driven BMI stratification via K-means and, within each stratum, determines the optimal testing time using decision-tree regression;third, it introduces a Gaussian mixture model for adaptive threshold labeling and, by jointly training Logistic/GBDT/Random Forest models with five-fold AUC–based selection, constructs a gestational-age–varying attainment-probability curve and, together with a Pareto frontier and sensitivity analysis, outputs an interpretable recommended T$*$. The findings indicate that this framework is robust to non-normality and heteroscedasticity, yields clear and reproducible grouping, is insensitive to perturbations of key parameters, and, while maintaining pass rates, enables earlier screening and reduces repeat blood draws, thereby enhancing the clinical operability and decision value of NIPT across BMI subpopulations. Despite the robust methodology, the current model lacks prospective validation across diverse clinical settings and does not fully integrate other confounding maternal physiological factors. Future work will focus on a multi-center prospective trial to externally validate the pipeline and extend the framework by incorporating additional biomarkers to construct a composite clinical utility index for more comprehensive prenatal guidance.

References

[1] Shi Weihui, Xu Chenming. Clinical value of noninvasive prenatal testing in the diagnosis of maternal obstetric complications and comorbidities[J]. Journal of Practical Obstetrics and Gynecology, 2025, 41(08):617-620.

[2] Jiang Liya, Lu Shaokan, Du Jiaen, Chen Kuoko, Zhang Sitian, Wu Shengjie, Liu Tianyu, Ye Jiahao. Development and application of noninvasive prenatal testing technology[J]. Clinical Medical Research and Practice, 2025, 10(23):191-194.

[3] Zhang Peng, Mo Weiying, Meng Minghui, Zhang Hongyan. Impact of noninvasive prenatal testing on detection of sex-chromosome aneuploidies and relevant ethical considerations[J]. Chinese New Clinical Medicine, 2025, 18(06):690-695.

[4] Zhang Zhenzhen, Chen Weiqing, Ding Ying. Value of combining noninvasive prenatal testing with theβ-subunit of human chorionic gonadotropin and pregnancy-associated plasma protein A for predicting fetal chromosomal abnormalities at 9–13+6 weeks of gestation[J]. Chinese Journal of Maternal and Child Health, 2025, 40(12):2165-2168.

[5] Yang Peng, Tang Kai, Zhang Juanling, Wang Wenjuan, Hu Xiaohong, Wang Qian. Value of ultrasound combined with NIPT in prenatal screening among advanced maternal age[J]. Laboratory Medicine and Clinical, 2024, 21(22):3348-3351+3358.

[6] Juanjuan Cheng. An empirical study on the relationship between university research and teaching:An analysis based on the Pearson correlation coefficient[J]. Chinese University Science and Technology, 2022, (10):46-52.

[7] Gong Jianwei, Zhang Linfeng. Application of generalized linear regression models in predicting postgraduate entrance examination scores[J]. Fujian Computer, 2022, 38(09):18-22.

[8] Aihemaiti Namaiti, Zeng Suiping, Tusun ai'i Tuerhong, Zeng Jian. Heterogeneity of urban form–thermal environment effects at the community scale and clustering partitioning:An integrated application of MGWR and K-means[J]. Advances in Geography, 2025, 44(01):199-210.

[9] Long Wenjia, Zhang Xiaofeng, Zhang Lian. A business-process clustering method based on K-means and the elbow rule[J]. Journal of Jianghan University (Natural Science Edition), 2020, 48(01):81-90.

[10] Wang Xinying, Song Guannan, Tang Qin. Factors associated with intraoperative hypoxemia during curative surgery for elderly lung cancer:Analysis based on logistic regression and decision-tree models[J]. General Practice Nursing, 2025, 23(11):2117-2121.

[11] Yang Mingyue. Day-ahead wind power forecasting based on a Gaussian mixture model and CNN–BiLSTM–Attention[J]. Electrical Applications, 2025, 44(05):86-96.

[12] Bao Shilong, Xu Qianqian, Yang Zhiyong, Hua Cong, Han Boyu, Cao Xiaochun, Huang Qingming. Efficient adversarial training for AUC optimization[J]. Journal of Computer Science, 2025, 48(07):1551-1571.