Integrating GAM and Survival Analysis: A Multi-Model Collaborative Framework for Optimizing NIPT Timing and Chromosomal Abnormality Determination in High-BMI Pregnant Women

Yu Yangyi Zhao Zhenyi Yi Yaxuan Zheng Shican

School of Materials Science and Engineering，Shenyang Aerospace University，Shenyang，110136;

Abstract： The application of non-invasive prenatal testing (NIPT) in pregnant women with high body mass index (BMI) faces challenges such as significant fluctuations in fetal cell-free DNA concentration, inconsistent testing timepoints, and complex chromosomal abnormality determination. Based on NIPT data from 1081 high-BMI pregnant women, this study constructed a collaborative analysis framework integrating generalized additive models (GAM), survival analysis, and multi-model fusion classification. First, GAM revealed nonlinear relationships between Y chromosome concentration and gestational age/BMI, achieving 82% model explanatory power. Second, Kaplan-Meier curves and Cox proportional hazards models stratified male fetus pregnancies by BMI and multifactorial risk factors. This identified optimal testing windows across groups that balanced high expected pass rates (≥90%) with low potential risks, while validating window stability. For female fetal chromosomal anomaly detection, a multi-output classification model integrating RandomForest, XGBoost, and LightGBM demonstrated superior performance in detecting abnormalities on chromosomes 13, 18, and 21 (most metrics exceeding 0.98). SHAP analysis further confirmed that the Z-score of the target chromosome serves as the core predictive feature. This study provides a comprehensive, interpretable clinical decision support solution for high-BMI pregnant women, spanning concentration prediction, timing optimization, and anomaly determination.

Keywords: Generalized additive model; Kaplan-Meier survival analysis; Cox proportional hazards model; LightGBM multi-output classification; SHAP interpretability analysis; Non-invasive prenatal testing

DOI： 10.69979/3041-0843.25.04.084

1 Introduction

Non-invasive prenatal testing (NIPT) has become a vital tool for screening fetal chromosomal aneuploidy (e.g., Down syndrome, Edwards syndrome, Patau syndrome) by analyzing fetal cell-free DNA fragments in maternal peripheral blood. Its high sensitivity and non-invasive nature enable widespread clinical application. However, the accuracy of NIPT is significantly constrained by fetal-specific DNA concentrations (e.g., Y-chromosome concentration in male fetuses) and the timing of testing. Clinical practice indicates that testing too early may result in sequencing failure, while testing too late may shorten the critical therapeutic window and increase risks for both mother and baby. Particularly for women with high BMI, individual physiological differences—such as blood circulation and fluid distribution—significantly impact the enrichment process of fetal cell-free DNA in maternal blood. This can delay the time required to achieve adequate Y-chromosome concentration or complicate the interpretation of chromosomal Z-scores, making a uniform testing timeline strategy difficult to apply. Therefore, establishing precise chromosomal abnormality prediction models through refined grouping based on key indicators like maternal BMI, optimizing NIPT timing for each group, and minimizing associated risks hold urgent clinical significance for enhancing prenatal screening efficacy in high-BMI populations. This study addresses this clinical challenge by integrating multiple statistical and machine learning models to construct a collaborative analysis framework. First, we employed Generalized Additive Models (GAM) to deeply analyze the nonlinear relationship between Y

chromosome concentration and gestational age/BMI. Second, we integrated survival analysis techniques (Kaplan-Meier and Cox models) to stratify risk for female fetuses and optimize testing timing. Finally, for male fetuses, we applied advanced ensemble learning models for chromosomal anomaly detection, enhanced by SHAP analysis to improve model interpretability.

2 Data Foundation and Preprocessing

This study utilized a dataset comprising NIPT test records from 1,081 high-BMI pregnant women. The data included key indicators such as gestational age, BMI, age, height, weight, Y chromosome concentration, Z-scores for each chromosome, GC content, and read counts. To ensure model reliability, rigorous data preprocessing was conducted: missing or invalid Y chromosome concentration records were removed; multiple test samples were averaged using weighted averaging; gestational age was uniformly converted to a continuous variable; continuous variables like BMI were standardized to eliminate dimensional effects; and multiple imputation methods were applied to handle missing values. Following preprocessing, data quality significantly improved, establishing a solid foundation for subsequent modeling. Descriptive statistical analysis results are presented in Table 1, illustrating the central tendency and dispersion of key variables.

Table 1 Descriptive Statistics Analysis

Variable	Y Chromosome Concentration	Gestational Age	Maternal BMI	Age
Sample Size	1081	1081	1081	1081
Mean	0.077	16.846	32.287	28.940
Standard Deviation	0.034	4.078	2.973	3.658
Minimum	0.010	11.000	20.703	21.000
25th percentile	0.051	13.286	30.196	27.000
Median	0.075	16.000	31.797	29.000
75th percentile	0.099	20.000	33.908	31.000
Maximum	0.234	29.000	46.875	43.000

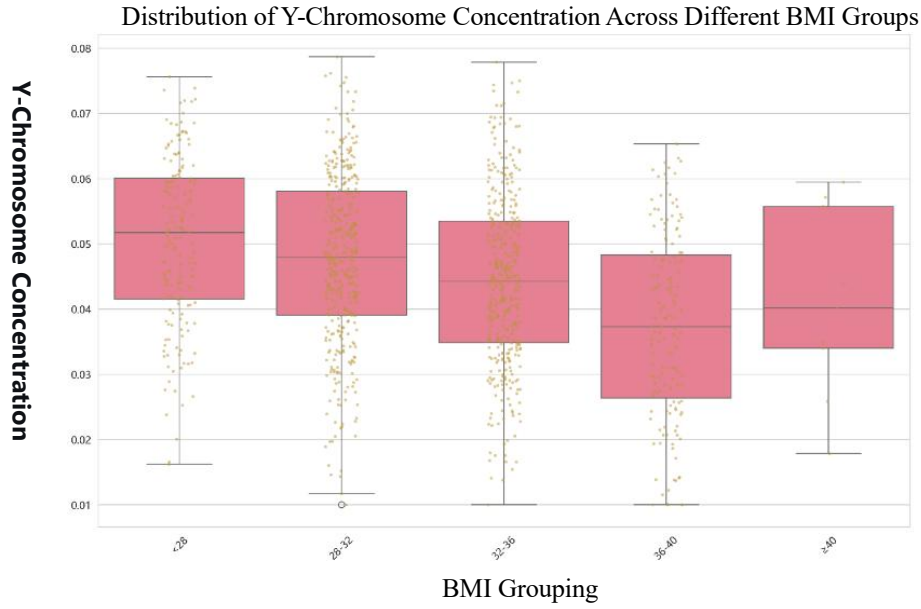

Figure 1. Distribution of Y-chromosome concentration across different BMI groups

Preliminary analysis indicates a distinct nonlinear relationship between Y chromosome concentration and variables such as gestational age and BMI. Furthermore, significant differences exist in Y chromosome concentration distribution across different BMI groups Figure 1, confirming the necessity of group-specific modeling.

3 Modeling the Nonlinear Relationship Between Y Chromosome Concentration and Gestational Age and BMI

To precisely characterize the complex relationship between Y chromosome concentration and gestational age and BMI, this study employed a Generalized Additive Model (GAM). The advantage of the GAM lies in its ability to flexibly capture nonlinear effects between variables through smoothing spline functions without requiring a predefined functional form. The model was constructed with Y chromosome concentration as the response variable and gestational age, BMI, and age as independent variables, accounting for the interaction between gestational age and BMI. A cubic B-spline basis was used for the smoothing function, and the model's smoothing parameters were optimized using the generalized cross-validation (GCV) criterion to balance model fit and complexity. Model results reveal a monotonically increasing nonlinear trend in Y chromosome concentration with gestational age, though the growth rate is not constant. BMI exhibits a significant nonlinear negative effect on Y chromosome concentration: pregnant women with higher BMI generally have lower Y chromosome concentrations, and the rate of increase with gestational age is slower. Age also demonstrated a significant influence. The model exhibited an overall explanatory power of 82%, indicating strong capture of data variability. Significance tests (F-tests) for each smoothing term yielded p-values below 0.001, confirming gestational age, BMI, and age as significant factors affecting Y chromosome concentration.

4 Optimizing NIPT Testing Timing for Male Fetus Based on Survival Analysis

Building upon the established patterns of Y chromosome concentration changes, this study further focused on optimizing the NIPT testing window for female fetuses. The core approach defined the first occurrence of Y chromosome concentration reaching or exceeding the 4% clinical threshold as an "event," treated gestational age as "time," and employed survival analysis techniques for modeling.

4.1 BMI Grouping Strategy

First, based on clinical practice and data distribution characteristics, the K-means clustering algorithm was used to divide maternal BMI into four clinically meaningful groups: normal group (BMI: [20.7, 28)), mildly overweight group (BMI: [28, 32)), moderately overweight group (BMI: [32, 36)), and severely overweight group (BMI ≥ 36). This grouping ensures homogeneity within groups and heterogeneity between groups, facilitating the development of differentiated strategies.

4.2 Kaplan-Meier Survival Curves and Optimal Timing Determination

For each BMI group, Kaplan-Meier survival curves were plotted for Y chromosome concentration failure to meet target levels. The optimal testing time point was determined based on the principle of "expected success rate ≥90% and potential risk <5%." By calculating the survival function, the earliest gestational week meeting this criterion was identified as the optimal time point for each group. Results revealed significant differences in optimal testing timing across BMI groups: 17.1 weeks for the normal group, 16.9 weeks for mildly overweight groups, and delayed to 24.7 weeks and 24.3 weeks for moderately and severely overweight groups, respectively. This clearly indicates that as BMI increases, pregnant women require later testing to ensure Y chromosome concentrations reach levels reliable for detection.

4.3 Robustness Analysis of Detection Error

To assess the impact of unavoidable random errors in actual testing on optimal timing estimates, sensitivity analysis was conducted using Bootstrap resampling techniques. Results showed that after accounting for detection error, the 95% confidence interval width for the optimal timing across all groups was only 0.1–0.2 weeks, with the high-BMI group exhibiting a slightly wider interval. This indicates that the optimal timing determined in this study possesses good stability

and reliability, but suggests that stricter quality control measures are needed for high-BMI pregnant women.

4.4 Multifactorial Cox Model and Risk Clustering

To comprehensively evaluate the combined effects of height, weight, age, and other factors, a Cox proportional hazards model was constructed. By calculating each pregnant woman's risk score and integrating her BMI distribution, K-means clustering was used to classify them into high- and low-risk groups. Defining the optimal timing as the "95th percentile of the target time distribution," the results indicated 24.8 weeks for the high-risk group and 17.7 weeks for the low-risk group. Bootstrap analysis further validated the robustness of this timing estimate. This approach provides a more refined individual risk assessment than simple BMI grouping.

Table 2 Optimal Screening Timing and Related Indicators for Each Risk Group

Risk Group	Optimal Timing (Weeks)	Expected Achievement Rate	Potential Risk Value	Sample Size	Average BMI
High-Risk Group	24.800	0.957	2.560	552	32.814
Low-risk group	17.700	0.952	1.140	529	31.737

Table 2 summarizes the optimized results based on Cox model risk stratification. This model framework successfully translates complex multifactorial information into concise clinical decision support, achieving the goal of minimizing potential risks by detecting at the earliest possible time point while ensuring testing accuracy.

5 Multi-model fusion decision for female fetal chromosomal abnormalities

For female fetuses, the absence of the Y chromosome as an internal reference necessitates relying on other indicators to determine chromosomal abnormalities. This study addresses trisomy anomalies on chromosomes 13, 18, and 21 (a multi-label classification problem) by constructing a robust multi-model fusion decision system.

5.1 Feature Engineering and Model Construction

Selected features include Z-scores of target chromosomes (13/18/21), X chromosome Z-score, GC content, proportion of valid reads, maternal BMI, and age. To address the class imbalance with few abnormal samples, the SMOTE algorithm was employed to generate synthetic data. Principal Component Analysis (PCA) retaining 95% variance was applied for dimensionality reduction while preserving key information. Based on this, RandomForest, XGBoost, and LightGBM models were trained separately, with cross-entropy loss and grid search used for hyperparameter optimization.

5.2 Model Performance and Interpretability

Model performance evaluation revealed that the ensemble model system achieved precision, recall, and F1-Score above 0.98 for detecting abnormalities in chromosomes 13 and 18. Accuracy and AUC both exceeded 0.99, demonstrating near-perfect discrimination capability. For detecting abnormalities on chromosome 21, the F1-Score was 0.86, AUC was 0.994, and macro-averaged AUC reached 0.986, also demonstrating excellent performance. To understand the model's decision-making basis and enhance clinical trustworthiness, SHAP values were employed for interpretability analysis. Results consistently indicate that the Z-score of the target chromosome itself is the most critical feature contributing to anomaly detection, far outweighing other features in importance. For instance, the Z-score of chromosome 18 plays a decisive role in identifying trisomy 18. While features like maternal BMI and age provide supplementary information, their contribution is relatively minor. This confirms the model's findings possess a robust biological foundation.

6 Conclusion

This study successfully established a multi-model collaborative framework integrating GAM, survival analysis, and ensemble learning, systematically addressing key clinical challenges in NIPT testing for high-BMI pregnant women. The GAM model revealed complex nonlinear relationships between Y chromosome concentration and gestational age/BMI, providing theoretical support for optimizing testing timing. Survival analysis enabled BMI-based and multifactorial risk

stratification for female fetuses, identifying optimal testing windows across groups that balance high detection rates with low risk, with robust performance validated through error analysis. For female fetal chromosomal anomaly detection, the constructed multi-model fusion system demonstrated high accuracy and robustness across trisomy 13, 18, and 21 screening. SHAP interpretability analysis further clarified the biological rationality of model decisions, enhancing clinical credibility. This framework organically integrates multiple modeling approaches, forming a comprehensive solution spanning mechanism exploration, timing recommendations, and anomaly detection. It significantly enhances the applicability and accuracy of NIPT in high-BMI populations. Despite limitations such as high dependence on data quality and computational complexity, future advancements can be driven through model lightweighting, incorporation of additional clinical indicators, and prospective validation to further promote its clinical translation. This study provides reliable theoretical support and practical tools for achieving personalized, precision NIPT testing in high-BMI pregnant women.

References

[1] Application of NIPT in prenatal screening for fetal chromosomal aneuploidy. Liu Li; Zhao Jinsong; Wang Dan; Xu Yan; Qi Chuan; You Junling; Yu Xiaochuan; Zhou Li. Chinese Journal of Eugenics and Genetics, 2019(11).
[2] Clinical Significance of NIPT in Detecting Rare Autosomal Aneuploidies. Hu Liang; Liu Jinxing; Wen Lijuan; Pei Yuanyuan; Luo Qi; Wei Fengxiang. Chinese Journal of Prenatal Diagnosis (Electronic Edition), 2022(03).
[3] Issues Related to NIPT for Prenatal Screening of Common Fetal Chromosomal Aneuploidies. Liu Juntao. Journal of Practical Obstetrics and Gynecology, 2018(11).
[4] Application of NIPT in Prenatal Screening for Fetal Chromosomal Aneuploidy. Zhi Diyuan. Chinese Journal of Occupational Health, 2021(05).
[5] Research Progress on NIPT Application in Diagnosing Fetal Aneuploidy Malformations. Teng Yue; Liu Dan; Huang Pu; Li Xu; Li Chunfang. Chinese Journal of Maternal and Child Health Research, 2016(02).

Co-first Author: Yu Yangyi, Zhao Zhenyi