World Journal of Information Technology
Print ISSN: 2959-9903
Online ISSN: 2959-9911
DOI: https://doi.org/10.61784/wjit3070

A FINE-TUNED STUDY ON OPTIMAL TIMING FOR NON-INVASIVE PRENATAL TESTING BASED ON GENERALIZED ADDITIVE MIXED MODELS AND SIMULATED ANNEALING OPTIMIZATION

XiangMeng Shu
School of Computer Science, Xi'an Shiyou University, Xi'an 710065, Shaanxi, China.

Abstract: Determining the optimal timing for non-invasive prenatal testing (NIPT) is critical, as it requires balancing the need for early detection with ensuring sufficient fetal DNA concentration for accuracy, particularly in male pregnancies. This study aims to optimize the NIPT window by analyzing the correlation between fetal Y chromosome concentration and maternal factors (gestational age and BMI) and proposing a stratified testing strategy. We first employed a Generalized Additive Mixed Model (GAMM) to capture complex nonlinear relationships, revealing a gradual increase in concentration with gestational age and a nonlinear inflection effect for BMI (with a slowdown after BMI~30). Subsequently, a hierarchical refinement strategy was implemented: pregnant women were clustered into homogeneous groups based on BMI and time-to-target concentration. A risk function quantifying both temporal and accuracy risks was then minimized using a simulated annealing algorithm to identify the optimal gestational week for each cluster. Results indicate that clusters with a BMI around 30 are suitable for early testing at 11 weeks, while high-BMI clusters require postponement to 24-25 weeks. This stratified approach significantly improved expected accuracy, with one group achieving 100%. The key innovations lie in using an interpretable GAMM for nonlinear analysis, data-driven clustering for population stratification, and a simulated annealing framework for balanced timing optimization.
Keywords: Chromosomal concentration; Generalized additive mixture model; Clustering; Simulated annealing; Optimal timing

1 INTRODUCTION

Non-invasive prenatal testing (NIPT) is a vital prenatal screening technique that evaluates fetal chromosomal abnormalities by analyzing fetal cell-free DNA in maternal peripheral blood. For male fetuses, achieving adequate chromosomal concentration is critical for testing accuracy. However, clinical practice reveals that fetal chromosome concentration is significantly influenced by factors such as maternal gestational age and body mass index (BMI). Particularly in women with high BMI, a dilution effect delays the attainment of optimal concentration levels. This creates an inherent clinical dilemma: balancing the urgency of "early detection to extend the therapeutic window" with the necessity of "ensuring sufficient concentration to guarantee accuracy." Previous studies often employed static, uniform standards to define testing windows or groupings, failing to adapt to the dynamic patterns of individual concentration changes and resulting in accuracy biases. This research addresses this core challenge by establishing a mathematical model to analyze the complex relationship between fetal chromosome concentration and maternal physiological characteristics, thereby enabling precise optimization of testing timepoints. The innovations of this section are: First, employing generalized additive mixture models to characterize the complex nonlinear relationships and inflection effects between chromosomal concentration, gestational age, and body mass index (BMI), providing interpretable marginal contributions and avoiding "black-box models"; Second, it overcomes the limitations of traditional empirical grouping by employing clustering methods to construct a hierarchical, refined BMI grouping scheme based on differences in target achievement time, thereby enhancing the consistency of strategies within each group[1-2]. Third, it constructs a risk quantification function and introduces a simulated annealing optimization algorithm to determine the optimal detection timing for each group while balancing "early detection" and "target achievement risk." The research protocol in this section followed these steps: First, concentration correlation analysis was conducted to establish models linking chromosome concentration to gestational age and BMI, followed by significance testing. Second, the optimal detection timing for male fetuses was optimized by implementing population stratification through clustering and solving for the optimal detection timing using the risk function and simulated annealing algorithm.

2 NONLINEAR MODELING AND ANALYSIS OF FETAL Y CHROMOSOME CONCENTRATION IN RELATION TO GESTATIONAL AGE AND BMI

2.1 Condition Assumptions

It is assumed that the relationship between gestational age and Y chromosome concentration is monotonically increasing, which conforms to general biological laws[3-5].

It is assumed that the impact of sudden events on the model can be ignored, such as sequencing failures or extreme outliers, which do not affect the main trend in the model to simplify the complexity of modeling.
It is assumed that there is a single generalized additive relationship between Y chromosome concentration and gestational age/BMI, avoiding unnecessary complexity while capturing the nonlinear relationships between variables.

2.2 Model Establishment

This problem aims to analyze the relationship between gestational age, BMI, and Y chromosome concentration. A Generalized Additive Mixed Model (GAMM) is adopted to establish this relationship model. GAMM is suitable for capturing nonlinear relationships between variables and can flexibly handle the impacts of different features; its form is:

$$Y_i=\alpha+s_1(t_i)+s_2(BMI_i)+\varepsilon_i \quad (1)$$

Where Y is the Y chromosome concentration (target variable), s_1 and s_2 are the smooth functions of gestational age and BMI respectively, α is the constant term, and ε is the error term.
The smooth function s(x) in the GAMM model is generally expressed as a basis function expansion:

$$s(x)=\sum_{k=1}^{K}\beta_k B_k(x) \quad (2)$$

Where: $B_k(x)$ is the basis function (commonly B-spline basis function or thin-plate spline basis function); β_k is the coefficient, obtained through model training and fitting; K is the number of basis functions (controlled by degrees of freedom or smoothness parameters). Substituting into the model gives:

$$Y_i=\alpha+\sum_{k=1}^{K_1}\beta_{1k}B_{1k}(t_i)+\sum_{k=1}^{K_2}\beta_{2k}B_{2k}(BMI_i)+\varepsilon_i \quad (3)$$

First, pandas and numpy are used for data cleaning: Z-score is adopted to handle outliers, median is used to fill missing values, and incorrect formats are converted into numerical formats. It is ensured that the Y chromosome concentration is within [0, 100]%, BMI is within [10, 50], and gestational age is within [10, 25] weeks. Secondly, the gestational age is converted from string format to numerical type, and other relevant features such as gestational age, BMI, and Y chromosome concentration (e.g., sequencing quality) are selected as independent variables.

2.3 Model Solution

Python is used for data processing, modeling, and result visualization: the pandas library for data processing, pyGAM for constructing the generalized additive model, scikit-learn for data standardization, and matplotlib for result visualization. The algorithm flow steps are as follows:
Step 1: Data Standardization and Normalization
Standardize or normalize the gestational age and BMI data to eliminate the impact of different dimensions, enabling comparison on the same scale.
Step 2: Regression Model Construction
A generalized additive mixed model is used for modeling to analyze the relationship between fetal Y chromosome concentration and maternal BMI/gestational age.
Step 3: Model Fitting
Fit the regression model, and construct smooth functions through the LinearGAM class to fit the relationship between variables and the target variable[6].
Step 4: Prediction
Use the trained model to predict the data and obtain the predicted values of Y chromosome concentration.

2.4 Result Analysis

2.4.1 Basic analysis

Using the scipy package in Python, the Pearson correlation coefficient r is calculated by substituting data to measure the strength and direction of the linear relationship between two variables:

$$r=\frac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i-\overline{y})^2}} \quad (4)$$

The Pearson correlation coefficient r measures the linear correlation between two continuous variables, with a value range of -1 to 1. When, it indicates a perfect positive correlation; when, it indicates a perfect negative correlation[7-8]; when, it indicates no linear relationship. The p-value is used to test whether the dependence between two continuous variables is significant, i.e., whether there is a statistically significant correlation. Under the condition of rejecting the null hypothesis, a p-value less than 0.05 indicates a significant correlation; a p-value greater than 0.05 indicates that the correlation may not be significant.

Calculations show that the correlation coefficient between gestational age and Y chromosome concentration is r=0.118 with, indicating a weak positive correlation between them—i.e., as gestational age increases, the Y chromosome concentration shows a slight upward trend.

The correlation coefficient between BMI and Y chromosome concentration is r=−0.155 with, indicating a weak negative correlation between them—i.e., as BMI increases, the Y chromosome concentration slightly decreases.

2.4.2 In-depth analysis

To reveal the regular relationship and biological mechanism between BMI/gestational age and fetal Y chromosome concentration, an in-depth analysis is conducted. During the GAMM fitting process, gestational age and BMI are used as variables, and spline functions are used to fit their impacts on Y chromosome concentration.

① Gestational Age vs. Y Chromosome Concentration

From the scatter plot of gestational age vs. Y chromosome concentration, the overall trend shows that as gestational age increases from 12 to 24 weeks, the distribution of points gradually moves upward. In the first trimester (12–14 weeks), most Y chromosome concentrations are concentrated between 0.05 and 0.1; after the second trimester (20 weeks onwards), more concentrations greater than 0.1 appear, even approaching 0.2. Regarding individual differences, at the same gestational age (e.g., 16 weeks), the Y chromosome concentrations of different pregnant women vary greatly, ranging from 0.04 to 0.15. This indicates that in addition to gestational age, BMI, GC content, and individual differences also affect the Y chromosome concentration. The overall upward trend is derived from the "upward shift of the point cloud" in the scatter plot and the correlation coefficient; the large individual differences are derived from the "wide distribution range of concentration points at the same gestational age".

The curve of gestational age vs. Y chromosome concentration is approximately monotonically increasing with a relatively uniform slope rather than severe fluctuations, indicating that the impact of gestational age on Y chromosome concentration is relatively gentle. The effect of gestational age on Y chromosome concentration shows a relatively stable increasing trend, which is consistent with the physiological law of gradual placental development and increased release of fetal cell-free DNA[9-10].

In summary, there is a significant nonlinear relationship between gestational age and Y chromosome concentration. As gestational age changes, the variation law of Y chromosome concentration is relatively complex, not a single linear trend—i.e., gestational age is a nonlinear factor affecting Y chromosome concentration.

② BMI vs. Y Chromosome Concentration

The scatter plot of BMI vs. Y chromosome concentration intuitively shows that the higher the BMI, the lower the overall Y chromosome concentration, which conforms to the medical mechanism of the "high BMI dilution effect". The curve of BMI vs. Y chromosome concentration shows an obvious inflection point: in the low BMI stage (<28), the Y chromosome concentration increases rapidly; after reaching the 30–32 range, the change slows down or even decreases, indicating a nonlinear inflection point effect. The impact of BMI has a nonlinear inflection point, suggesting that in the high BMI population, the dilution effect of maternal blood on fetal DNA is more significant, thereby delaying the time when the concentration meets the standard. This finding indicates that a unified detection time point may be unfair to pregnant women of different body types, and BMI-stratified detection is more clinically valuable.

In summary, from the scatter plot and curve, BMI and Y chromosome concentration are not simply linearly related but show a complex nonlinear correlation: the direction and degree of the impact of different BMI intervals on Y chromosome concentration are different, indicating that BMI is a nonlinear factor affecting Y chromosome concentration. These nonlinear relationships verify that GAMM can effectively capture the complex impact relationships between gestational age, BMI, and Y chromosome concentration. A residual normality test plot is drawn.

2.4.4 Residual analysis

From the residual normality test plot, although some blue scatter points deviate from the red reference line, they generally distribute around this line, indicating that the residuals basically meet the characteristics of a normal distribution, satisfying the assumption of residual normality required by many statistical models.

Regarding the relationship between residuals and predicted values: in the residual analysis plot, the residual points are relatively scattered within the range of predicted values without obvious trends or patterns, indicating that the residuals are random. The prediction error of the model is relatively stable at different predicted value levels, and the model's fitting effect is good in terms of residual randomness.

2.5 Model Verification

2.5.1 Significance test

The LRT test is used to determine whether the impacts of gestational age and BMI on Y chromosome concentration are significant.

Note: In the built-in significance test of the model (GAMM), the p-values of both gestational age and BMI are displayed as 1.11e−16. It should be noted that this value is the minimum positive lower limit that can be represented by computer floating-point numbers, indicating that both have a highly significant impact on Y chromosome concentration. Due to the large sample size or strong variable effects, the p-values have approached the theoretical zero value, so it is impossible to further distinguish their specific magnitude differences. The equal p-values here do not mean that the effect sizes of the two are the same, only that their impacts are both statistically significant.

As can be clearly seen from the table below, the p-values of both gestational age and BMI are less than 0.05, indicating a significant correlation. Significance Test Results is shown in table 1.

Table 1 Significance Test Results

Variable	p-value	Significance Judgment
Gestational Age	1.11×10^{-16}	Significant
BMI	1.11×10^{-16}	Significant

2.5.2 Result evaluation

The Pearson correlation coefficient r^2 is used to measure the goodness of fit of the regression model to the data:

$$r^2=1-\frac{SSR}{SST} \tag{5}$$

Where: $SSR=\sum (Y_i-\hat{Y}_i)^2$ (sum of squared residuals), $SST=\sum (Y_i-\overline{Y})^2$ (total sum of squares), Y_i is the actual value, $\hat{Y}_i$ is the predicted value, and $\overline{Y}$ is the mean of the observed values.

The value range of R^2 is between 0 and 1. The closer R^2 is to 1, the better the model fits, indicating that the model explains most of the fluctuations of the dependent variable; the closer R^2 is to 0, the worse the model fits, indicating that the model can hardly explain the fluctuations of the dependent variable. The calculated coefficient of determination, proving that the model fits well.

2.5.3 Residual Normality Test

Calculate the residual of each observation point, i.e., the difference between the predicted value and the actual value of the model:

$$e_i=Y_i-\hat{Y}_i \tag{6}$$

Observe whether the quantiles of the residuals are close to the quantiles of the standard normal distribution through the residual normality test plot. If the points are distributed along a straight line, the residuals are close to a normal distribution.

Observe whether the residuals are randomly distributed around 0 without systematic trends through the residual scatter plot. The degree of dispersion of the residuals should be roughly the same, and there should be no funnel shape (i.e., the residuals become more scattered as the predicted value increases). The results of the residual plot test show that the residuals are roughly randomly distributed, indicating that the model's fitting effect is good. The QQ plot shows that the residuals basically conform to a normal distribution, further verifying the fitting quality of the model.

3 OPTIMIZING THE TIMING OF MALE FETAL NIPT DETECTION USING K-MEANS CLUSTERING AND SIMULATED ANNEALING ALGORITHMS

3.1 Condition Assumptions

3.1.1 Risk quantification assumption

It is assumed that potential risks can be obtained by weighting time risk and accuracy risk, i.e.,

$$R=\alpha\times time\_risk+\beta\times(1-p) \tag{7}$$

Where, time_risk is the time risk, p is the predicted compliance probability, and $(1-p)$ is the accuracy risk.

Its rationality lies in comprehensively considering two key factors: early screening in the first trimester and the accuracy of test results. The weights can be adjusted to reflect the importance of different factors. It not only inherits the precondition of "detection time point" but also supplements the key dimension of test result quality. At the same time, the weights can be flexibly adjusted to reflect the priority of the two major factors in different scenarios, forming a core logical closed loop from "operational rules to risk quantification".

3.1.2 Clustering stability assumption

This assumption aims to ensure the reliability of data classification during risk quantification, i.e., it is assumed that K-means clustering can converge to stable results under given parameters. This assumption is based on the inherent nature of the K-means algorithm: under reasonable parameter settings, the algorithm can achieve stable clustering centers after a limited number of iterations, thereby providing reliable algorithmic support for the stratified analysis of data related to time risk and accuracy risk, and ultimately ensuring the validity of the results of the entire risk quantification system.

3.1.3 Detection time point discreteness assumption

From the perspective of the implementability of actual detection operations, the detection time point discreteness assumption is proposed, i.e., the detection time point can only take integer gestational weeks. This assumption is consistent with the actual practice in medical scenarios where detections are usually arranged by whole weeks, providing a "time dimension foundation" consistent with real operations for all subsequent risk analysis and data processing.

3.2 Model Establishment and Solution

3.2.1 BMI–compliance time model

To reveal the relationship between maternal BMI and the compliance time of fetal cell-free DNA, and automatically discover homogeneous populations in the "BMI-compliance time" space to reduce intra-group heterogeneity and

improve the consistency of strategies within groups. Before establishing the model, K-means is first used for stratification in the two-dimensional space to lay the foundation for subsequent group-specific recommendation of detection time points and risk assessment; second, male fetus maternal data is screened from the processed data, and relevant data such as BMI and Y chromosome concentration compliance time are extracted; at the same time, K-means clustering parameters are initialized, such as the initial position of clustering centers and the maximum number of iterations.

Model Establishment

① Data Preprocessing

To obtain a clean, dimensionless, and comparable dataset, data preprocessing is performed: unify gestational age into decimal weeks, convert FF to proportional form, remove outliers using IQR, unify units to reduce systematic bias, eliminate caliber differences, and obtain 1047 samples.

② Standardization

Z-score standardization is performed on BMI and compliance time variables to eliminate the dimensional difference between BMI and τ and make the distance measurement fair:

$$\tilde{x}=\frac{(x-\mu)}{\sigma} \tag{8}$$

③ Initialization of Clustering Centers

K-means++ is used to select new centers with a probability proportional to the distance from existing centers, which theoretically provides better initial coverage, thereby improving convergence speed and globality and reducing inferior solutions caused by initialization.

④ K-means Clustering

Define the distance between sample point x_i and clustering center c_j as:

$$d(x_i,c_j)=\sqrt{(x_{i1}-c_{j1})^2+(x_{i2}-c_{j2})^2} \tag{9}$$

Where x_{i1} and x_{i2} are the BMI and Y chromosome concentration compliance time of sample x_i respectively, and c_{j1} and c_{j2} are the corresponding coordinates of clustering center c_j. Continuously adjust the clustering centers to minimize $\sum_{j=1}^{k}\sum_{x_i\in cluster_j} d(x_i,c_j)^2$.

Calculate the Euclidean distance between sample points and clustering centers to measure the similarity between "samples and centers" and provide criteria for assignment:

$$d_{ij}=\sqrt{(BMI_i-\mu_{BMI,j})^2+(T_i-\mu_{T,j})^2} \tag{10}$$

⑤ Sample Belonging Judgment

Classify each sample into the nearest cluster according to the distance between the sample and the clustering center.

⑥ Update Clustering Centers

In the standardized space, take the sample mean of each cluster to make the center represent the "centroid" of the current cluster and continue to reduce the within-class sum of squares:

$$\mu_{BMI,j}=\frac{1}{n_j}\sum_{x_i\in cluster_j} BMI_i,\ \mu_{T,j}=\frac{1}{n_j}\sum_{x_i\in cluster_j} T_i \tag{11}$$

⑦ Iterative Solution

Repeat the above process until the objective function converges, monitor the objective function (within-class sum of squares) to obtain stable cluster division and centers, and stop improving the optimization objective:

$$J=\sum_{j=1}^{k}\sum_{x_i\in cluster_j} \|x_i-C_j\|^2 \tag{12}$$

Where, $C_j=(\mu_{BMI,j},\mu_{\tau,j})$.

Model Solution

Python simulation is adopted, with the number of clusters $K=4$, $n_{init}=10$, and random seed 42. The BMI and compliance time ranges of the 4 groups are as table 2:

Table 2 BMI and Compliance Time Table

Cluster	BMI Range	BMI Mean	Compliance Time Mean (Weeks)	Sample Size
0	27.0 – 33.4	30.9	21.9	224
1	32.0 – 39.3	33.8	14.6	327
2	26.6 – 31.9	30.0	14.3	367
3	33.5 – 39.4	36.0	21.9	129

Result Analysis

① Basic Analysis

Pregnant women with lower BMI (e.g., in the range of 26 – 30) can mostly meet the threshold around 11 – 14 weeks; pregnant women with higher BMI (>33) mostly need 20 weeks or even later to meet the threshold. This shows an obvious trend of "the higher the BMI, the longer the compliance time" in visualization, i.e., BMI is significantly positively correlated with compliance time, and pregnant women with high BMI are more likely to have delayed compliance.

② In-depth Analysis

Further analyze the clustering results to obtain the compliance detection time:

Table 3 BMI Mean and Compliance Detection Time Table

Cluster	BMI Mean	Compliance Time Mean (Weeks)
2	30.0	14.3
1	33.8	14.6
0	30.9	21.9
3	36.0	21.9

BMI Mean and Compliance Detection Time Table are shown in table 3.

The clustering results show that BMI≈30 is the critical point. Cluster 2 (BMI≈30) meets the standard at 11-14 weeks, indicating that it is suitable for early detection; Cluster 3 (high BMI group) generally meets the standard after 20 weeks and requires delayed detection.

③ Direct Response to the Problem

Based on modeling the relationship between BMI and Y chromosome compliance time, the model proposes a stratified and refined BMI grouping scheme, effectively avoiding the drawbacks of large intra-group differences and lack of clinical significance in boundary for traditional empirical grouping. Through quantitative grouping results, pregnant women in different BMI groups have higher consistency and predictability in compliance time, thereby providing a scientific and reliable modeling basis for the establishment of risk functions and the optimization of optimal detection time points.

Model Verification

To verify the model assumptions, two methods are used to test the model performance. The first is to use the silhouette coefficient to measure the compactness within clusters and separation between clusters, with a value range of [-1, 1], and the closer to 1, the better the clustering effect; the second is the Calinski-Harabasz (CH) index to measure the ratio of between-class variance to within-class variance, and the larger the value, the more significant the clustering effect.

The silhouette coefficient is:

$$s(i)=\frac{b(i)-a(i)}{\max\{a(i),b(i)\}} \tag{13}$$

Where: s(i) is the silhouette coefficient of a single sample; a(i) is the average distance from sample i to other samples in the same cluster, i.e., within-class compactness; b(i) is the average distance from sample i to samples in the nearest neighboring cluster, i.e., between-class separation. s(i) takes values in the interval [-1, 1].

The average silhouette coefficient is:

$$SC=\frac{1}{n}\sum_{i} s(i) \tag{14}$$

The closer SC is to 1: samples within clusters are compact and different clusters are well separated;

The closer SC is to 0: samples may be on the boundary, and the clustering effect is general;

SC is negative: the clustering is unreasonable, and samples are misclassified into clusters.

The result is an average silhouette coefficient, ensuring "no serious clustering errors and basic distinguishability between clusters".

The CH index is defined as follows:

$$CH=\frac{Tr(B_k)/(k-1)}{Tr(W_k)/(n-k)} \tag{15}$$

Where: B_k is the between-class scatter matrix, i.e., the variance between cluster centers; W_k is the within-class scatter matrix, i.e., the within-class compactness; n is the number of samples; k is the number of clusters.

A larger k value indicates significant differences between clusters (well separated) and compact within clusters (well clustered); there is no fixed threshold, and it is mainly used for comparison between different k values.

CH≈295 (optimal among candidate k values) ensures "significant differences between clusters and sufficient compactness within clusters". Combined, they support the conclusion of "reasonable clustering structure".

3.2.2 Risk optimization model for optimal detection time point

Model Establishment

To balance "early detection timing" and "high detection accuracy", a simulated annealing optimization model based on risk function is constructed, allowing the algorithm to accept worse solutions in the early stage to avoid falling into local optima and approach the global optimum as the temperature gradually decreases. The objective of this sub-problem is to determine the optimal NIPT detection gestational age for different BMI groups. Simulated annealing optimization is adopted to find the detection time point under the meaning of minimizing the risk function.
Risk quantification formula:

$$R(t)=\alpha\cdot R_{time}(t)+\beta\cdot(1-P_{FF}(t)) \tag{16}$$

Where: $R_{time}(t)$ is the time risk: assign 1 if $t\leq 12$; 3 if $13\leq t\leq 27$; 10 if $t>27$;
Accuracy risk:, where $P_{FF}(t)$ is obtained by the interpolation function; the weights are set as, $\beta=0.4$.
The purpose of optimization is to gradually reduce the probability of accepting worse solutions, making the search transition from "global search" to "local refinement", and finally obtain:

$$T_{k+1}=\gamma T_k,\ \gamma\in(0,1) \tag{17}$$

Model Solution
To realize the search for the optimal detection time point, this problem first performs a smooth approximation of the gestational age-compliance rate function through scipy.interpolate based on the Python environment, and then introduces the simulated annealing algorithm to iteratively solve the minimum value of the risk function with the idea of global optimization.
Step 1: Construct the Gestational Age–Compliance Rate Function P(t)
We have original data on whether the fetal DNA of each pregnant woman meets the standard (e.g., FF≥4% is recorded as compliant) at a certain gestational age g^i. The problem is that the data points we have are discrete, but we need to use the "compliance rate" in the risk function R(t) at any gestational age t. Therefore, it is necessary to convert the "discrete empirical proportion" into a "continuous smooth function". Summarize the gestational age by integer weeks: n_k : the total number of people tested at gestational age t_k; m_k: the number of compliant people among them; then the empirical compliance rate is:

$$p_k=\frac{m_k}{n_k} \tag{18}$$

$$p(t)=p_0+\frac{p_1-p_0}{t_1-t_0}\cdot(t-t_0) \tag{19}$$

Step 2: Initialization
Initialize, and calculate:

$$t'=\begin{cases}10, & \text{if } t_{best}+\Delta t<10\\ t_{best}+\Delta t, & \text{if } 10\leq t_{best}+\Delta t\leq 25\\ 25, & \text{if } t_{best}+\Delta t>25\end{cases} \tag{20}$$

Step 3: Random Search
Perform symmetric neighborhood random search on the discrete domain D to ensure state reachability and ergodicity. Randomly perturb $\Delta t=\pm 1$ among candidate solutions.
Step 4: Metropolis Criterion
Using the Metropolis criterion, regard the risk R(t) as "energy", and let the Markov chain have the Boltzmann distribution $\pi(t)\propto\exp(-R(t)/T)$ as the equilibrium distribution at temperature T. This can not only move towards lower risks but also jump out of local optima with a certain probability. If $\Delta R\leq 0$ (the new solution is better), accept it unconditionally: $t=t'$; if $\Delta R>0$ (the new solution is worse), accept the new solution with probability:

$$P=\exp\left(-\frac{\Delta R}{T}\right) \tag{21}$$

That is, let $t=t'$ with probability p. Maintain exploration in the high-temperature stage (high probability of accepting worse solutions) and tend to greedy convergence in the low-temperature stage (probability tends to be small), thereby improving the possibility of obtaining the global optimum.
Step 5: Decreasing Temperature Control
Control the transition of "exploration → exploitation" through decreasing temperature. Theoretically, slow cooling (such as logarithmic cooling) can ensure convergence to the global optimum; geometric cooling is often used in engineering to balance efficiency and effect:

$$T_{k+1}=\frac{T_0}{1+ck}\ \text{or}\ T_{k+1}=\frac{T_0}{\log(2+k)} \tag{22}$$

The temperature gradually decreases until convergence.

Table 4 BMI Mean and Optimal Detection Time Table

Cluster	Average BMI	Optimal Detection Time (Weeks)	Minimum Risk Value	Sample Size
0	30.9	25	1.848	224
1	33.8	16	1.844	327
2	30.0	11	0.600	367

3	36.0	24	1.901	129

Result Analysis

First, conduct a basic analysis of Table 4. Cluster 2 (medium-low BMI) has the optimal detection time of 11 weeks with the minimum risk (0.600); Clusters 0 and 3 (high BMI) need to be delayed until 24 – 25 weeks. Further in-depth analysis shows that sensitivity analysis indicates that the risk function is more sensitive to the compliance rate parameter (weight contribution ≈ 0.65), indicating that FF compliance is the core driving factor.

Model Verification

To verify the reliability of the optimization results, the "compliance rate within the optimal detection gestational age" is adopted. If the FF compliance proportion is significantly improved within the optimal detection gestational age, the model is effective. Accuracy verification (calculate the FF compliance proportion at the optimal detection time):

Table 5 Optimal Detection Time and Prediction Accuracy Table

Cluster	Optimal Detection Time (Weeks)	Expected Accuracy	Compliant Samples / Total Samples
0	25	87.91%	189 / 215
1	16	89.11%	221 / 248
2	11	100.00%	10 / 10
3	24	74.77%	80 / 107

Optimal Detection Time and Prediction Accuracy Table are shown in table 5. Cluster 2 can achieve 100% compliance when tested at 11 weeks, and the accuracy of Cluster 0 at 25 weeks is close to 88%. This indicates that testing according to the time point recommended by the model has a higher success rate, which is consistent with the medical mechanism —i.e., pregnant women with high BMI have relatively diluted fetal cell-free DNA in plasma, so FF compliance is later; pregnant women with medium-low BMI have earlier FF compliance. This indicates that the model is not only mathematically effective but also reasonable in medical interpretation. Therefore, this model can directly guide clinical practice, indicating that the optimization results have high practicality and reliability.

4 CONCLUSIONS

This study successfully established a modeling framework that clarifies the nonlinear relationships between fetal Y chromosome concentration, gestational age, and maternal BMI, and translates these insights into an optimized, stratified strategy for determining the optimal timing of NIPT in male pregnancies. The Generalized Additive Mixed Model (GAMM) revealed a monotonically increasing effect of gestational age and a nonlinear inflection effect of BMI. Leveraging these findings, a data-driven approach combining K-means clustering and simulated annealing optimization effectively identified distinct optimal testing windows for different BMI groups (e.g., 11 weeks for a BMI of ~30 and 24-25 weeks for higher BMIs), significantly enhancing detection accuracy. This work provides a robust and personalized decision-support tool for clinical prenatal screening.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] İbrahim Taşkum, Selcan Sınacı, Seyhun Sucu. Evaluating the reliability and clinical utility of artificial intelligence in first trimester prenatal screening and noninvasive prenatal testing. Scientific Reports, 2025, 15(1): 41331 - 41331.

[2] Rich M, Jan A S, Fraser C, et al. Outcome data for non-invasive prenatal testing suggestive of an atypical sex chromosome abnormality of fetal/placental origin. Journal of genetic counseling, 2025, 34(6): e70138.

[3] Leverington J, Keegan A, Azmanov N D, et al. Clinical Implementation of RHD NIPT in a Tertiary Obstetric Centre in Western Australia. The Australian & New Zealand journal of obstetrics & gynaecology, 2025.

[4] Godino L, Nardi E, Lanzoni G, et al. Attitudes, awareness and experience of Italian women undergoing non-invasive prenatal testing (NIPT): a nationwide cross-sectional study. BMJ open, 2025, 15(11): e102811.

[5] Zhang B, Chen X, Xi S, et al. Fetal fractions mediate the association between total cell-free DNA and preeclampsia risk in a non-invasive prenatal testing cohort. Human Genomics, 2025, 19(1): 133-133.

[6] Zemanick T E, Putra M, Elfman H, et al. Letter to the editor: False reassurance following single gene non-invasive prenatal testing for cystic fibrosis. Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society, 2025.

[7] Schuurman P V L, Koning D J H, Meier E, et al. Clinical and economic impact of genome-wide non-invasive prenatal testing (NIPT) as a first-tier screening method compared to targeted NIPT and first-trimester combined testing: A modeling study. PLoS medicine, 2025, 22(11): e1004790.

[8] Perrot A, Smart B H, Klaiman N T, et al. Decision-making for termination of pregnancy following non-invasive prenatal testing: a qualitative exploration of french, english and German healthcare professionals' perceptions and concerns. Reproductive health, 2025, 22(1): 216.

[9] Galeva S, Stoilov B, Uchikova E . Challenges and clinical implications of discordant non-invasive prenatal testing results: insights from two case studies. Folia medica, 2025, 67(5).

[10] Resta C, Xiong R, Sturrock S, et al. Non-invasive prenatal testing for the diagnosis of sickle cell disease in high-risk pregnancies: A systematic review and statistical summary of the current literature. European journal of obstetrics, gynecology, and reproductive biology, 2025, 316114799.