Mathematical Modeling and Algorithm Application
ISSN: 3006-0842 | Vol. 7, No. 2, 2025

A Data-Driven Event-Time Optimization Model Based on Individualized Risk Segmentation

Jie Ma*

Tianjin University of Science and Technology College of Electronic Information and Automation Department of Mathematics and Applied Mathematics, Tianjin, China

* **Corresponding author**: Jie Ma (Email: 1755646081@qq.com)

Abstract: Event-driven decision optimization has become a critical research area in computational modeling, especially in systems where outcomes depend on threshold-based signals and individual feature heterogeneity. This study proposes a data-driven event-time optimization model that determines the optimal decision timing under uncertain and feature-dependent conditions. The framework considers individual-level variations—such as body-mass-related indicators—as covariates influencing the time required for a measurable signal to reach a predefined threshold. A feature-based grouping strategy is developed to cluster individuals with similar dynamic patterns, assigning each group an optimal initial decision point. For cases where the threshold is not reached, a secondary decision (re-evaluation) mechanism is introduced to minimize overall cost and risk. By integrating risk segmentation functions, survival-type event-time modeling, and cost-sensitive optimization, the proposed model achieves improved accuracy, lower uncertainty, and higher efficiency in threshold-based decision systems. This framework not only enhances reliability within its biomedical origin but also provides a transferable methodology for broader applications in computational intelligence, predictive maintenance, and big-data-driven decision optimization.

Keywords: Event-time modeling; Risk-aware optimization; Data-driven grouping; Threshold-based decision system; Cost-sensitive learning

1. Introduction

Event-time prediction and individualized decision optimization have become increasingly important in modern data-driven systems, where outcomes depend on both dynamic signals and heterogeneous individual characteristics. Although this research originates from biomedical data, the methodological framework extends to general event-based modeling tasks characterized by threshold-driven outcomes, uncertainty, and cost–risk trade-offs. Non-invasive prenatal testing (NIPT) provides a representative example of such a system, where the objective is to identify the optimal decision timing under individual-specific features and noisy measurement processes. Despite significant progress in improving detection accuracy, the reliability of prediction models remains constrained by the variability of signal concentration dynamics and feature heterogeneity [1]. When the observable signal—such as a chromosome-related biomarker—fails to exceed a specified threshold, the system's output cannot be confirmed with high confidence. Moreover, if the decision (testing) is made too early or too late, the overall risk and cost of the process increase. Thus, determining an optimal event-time decision that adapts to individual features (e.g., body composition or other covariates) remains a non-trivial optimization challenge in both computational and applied domains.

Previous studies have partially examined how feature variations, such as body-mass-related indicators, influence the time required for a signal to reach its critical threshold. However, the quantitative relationship between individual covariates and threshold-achievement time has not been fully modeled, and the joint effect of measurement noise and feature heterogeneity on prediction reliability is still insufficiently explored. In addition, a systematic framework to balance initial decision timing, repeated evaluation, and cost–risk constraints has not yet been established [2–4].

To address these gaps, this study proposes a data-driven event-time optimization model based on individualized feature grouping and risk segmentation. Instead of relying solely on domain-specific assumptions, the proposed framework constructs a generalized optimization process: individuals are clustered into feature-based groups, and a unified initial decision time is assigned to each group through risk-aware optimization. For cases in which the threshold is not met, a dynamic re-decision mechanism is introduced to minimize expected cost and cumulative risk. By integrating segmented risk functions, survival-type event-time modeling, and cost-sensitive optimization, this approach quantifies the influence of individual covariates on event-time distribution and provides an interpretable, computationally efficient solution for threshold-based decision problems.

The significance of this study lies in its ability to generalize a domain-specific timing problem into a scalable computational framework. The proposed model demonstrates how individualized risk modeling and data-driven optimization can jointly reduce uncertainty and improve decision reliability across heterogeneous systems. Beyond its biomedical origin, this framework can be extended to other applications such as predictive maintenance, risk monitoring, or scheduling under uncertainty, offering a robust paradigm for computational modeling and big-data-driven decision optimization.

2. Optimal NIPT Timing Allocation Based on BMI

In this study, we propose a BMI-based scientific grouping method for pregnant women carrying male fetuses. By dividing the BMI values into a few groups, we aim to determine the optimal NIPT testing time for each group to

reduce potential risks during the testing process. Specifically, we define a unified initial NIPT testing week for each BMI group to ensure that pregnant women undergo testing at the most appropriate time, thereby enhancing the accuracy and success rate of detection [5-6]. If the initial test does not provide a clear result, we plan a retesting process based on predefined standards.

The core objective of this model is to balance the timing and costs of the initial test and retesting effectively. We will analyze how BMI differences impact the time required for NIPT to meet the established threshold and examine how errors during testing may affect the accuracy of the final results. By considering these factors, this study aims to offer an optimized testing strategy for clinical practice that minimizes potential risks, controls errors, and optimizes the additional costs and trade-offs involved in the retesting process.

2.1. The earliest compliance time (event time)

In this study, we define $Y_i(D)$ as the Y chromosome concentration of the i-th pregnant woman carrying a male fetus at gestational week D. To accurately reflect the clinical threshold, we use a 4% Y chromosome concentration, after error correction, as the threshold for meeting the standard. The "earliest threshold achievement time," A_i, is defined as the minimum gestational week D at which $Y_i(D) \geq 4\%$, meaning that the Y chromosome concentration first reaches the clinical standard at this gestational week. If, during the entire observation period, the Y chromosome concentration of the pregnant woman never reaches 4%, A_i is considered "right-censored data." This data type aligns with survival/event-time analysis frameworks and will be used for subsequent statistical modeling and analysis.

2.2. Risk and Segmentation Cost

In clinical practice, it is generally believed that the later the confirmation of fetal health status, the higher the potential risks. To quantify and describe this risk, this study employs a "piecewise constant function" to approximate the risk model. Specifically, the risk function $\varphi(D)$ is divided into three intervals based on the gestational week D:

φ_1 (lowest risk): When $D \leq 12$ weeks, representing the early stage;

φ_2 (increasing risk): When $13 \leq D \leq 27$ weeks, representing the mid-stage;

φ_3 (very high risk): When $D \geq 28$ weeks, representing the late stage.

By segmenting in this manner, we can effectively assess the potential risks at different stages of pregnancy, providing a basis for developing an optimized NIPT testing strategy.

2.3. Decision Point and Confirmation Time

In the execution strategy, we first set a unified initial testing gestational week D_g for each BMI group g. If the initial test does not yield a clear result (including cases where the Y chromosome concentration does not meet the threshold or sequencing fails), a retest will be conducted after an interval of X weeks. The individual's confirmation time, which is the moment when a reliable result is obtained, is calculated according to the following rule:

$$B(D_g) = \begin{cases} D_g, & \text{if } A_i \leq D_g \\ \max(A_i, D_g + X), & \text{if } A_i > D_g \end{cases}. \quad (1)$$

If the Y chromosome concentration meets the threshold and the test is successful at gestational week o_g, the confirmation time is o_g. If the Y chromosome concentration does not meet the threshold or the test fails at gestational week o_g, the confirmation time is the later of A_i (the earliest time the threshold is reached) and $o_g + X$ (the retest gestational week).

In general, if the initial test successfully meets the threshold, the result is confirmed immediately. If the initial test fails, the confirmation time will be the later of the two time points—either when the fetal Y chromosome concentration naturally reaches the threshold or the designated retest time. This strategy ensures a reasonable balance between the timing of the initial test and retesting, while reducing potential risks.

2.4. Optimization objective

In this study, the objective is to select the appropriate initial testing gestational week (o_g) such that the "expected risk" within a given BMI group (g) is minimized. The expected risk not only considers the accuracy of the test results but also incorporates the additional costs incurred from test failures (φ), where φ represents a weighted combination of time costs, economic costs, and maternal anxiety.

The optimization objective can be mathematically expressed as:

$$\text{Minimize}\left[E(\varphi(B_i(o_g)) \mid i \in g) + \varphi \times P(\text{Unconfirmed on first test} \mid i \in g)\right] \quad (2)$$

where $E(\varphi(B_i(o_g)) \mid i \in g)$ represents the expected risk for individuals within group g, i.e., the anticipated risk for each pregnant woman at the time of the initial test. $P(\text{Unconfirmed on first test} \mid i \in g)$ denotes the probability that the first test within group g yields an unconfirmed result. φ is a tuning parameter used to balance the different risks and costs.

The key point is that the value of φ directly influences the decision-making tendency, determining whether to choose "earlier retesting to reduce waiting time" or "delaying the test to improve the probability of a successful first test." Through this optimization objective, the study aims to rationally balance the costs and risks associated with retesting while minimizing potential risks, thereby selecting the most suitable timing for the initial test. The entire process of establishing the CRB-NIPT model is shown in Algorithm 1.

Algorithm 1. Training and Calibration of the CRB-NIPT Model

Input: Historical test data, Significance level α; retest interval Δ; candidate BMI group numbers K_{grid}; gestational week grid T_{grid}; risk function $\varphi(D)$; cost weight λ.

Output: BMI group set $\mathrm{G} = g$; per-group policy $\pi_g = (D_g^*, \Delta)$;

group-wise success model $p_g^{\text{succ}}(t,b)$; reach-time CDF $H_g(t)$.

1. Noise and threshold calibration

Estimate noise variance σ from short-interval retests.

Compute the effective threshold $e_4 = 0.04 + z_{1-\alpha}\sigma$.

2. Event-time construction

For each subject $i = 1, \ldots, N$:

(a) Fit a monotone smoother $\hat{y}_i(t)$ to (t_{ij}, y_{ij}) ensuring non-decreasing trend.

(b) Determine the earliest reach time $A_i = \inf t : \hat{y}_i(t) \geq e_4$:
If the threshold is never reached, mark the observation as right-censored $X_i = 1$.

3. BMI-driven grouping

Train a CART tree using $(\text{BMI} i \to (A_i, X_i))$.

Select the optimal number of groups $K \in K_{\text{grid}}$ by cross-validated expected risk reduction.

Output BMI intervals G.

4. Within-group modeling

For each group $g \in \mathrm{G}$:

(a) Estimate $H_g(t)$ and $S_g(t) = 1 - H_g(t)$ using Kaplan–Meier.

(b) Fit a logistic model $p^{\text{succ}} g(t,b)$ to predict sequencing success.

(c) For each $t \in T_{\text{grid}}$:

i. Compute the one-shot confirmation probability
$\theta_g(t) = p^{\text{succ}} g(t, \bar{b}_g) \cdot H_g(t)$.

ii. Calculate expected confirmation time
$E[T\text{conf} \mid t] = t \cdot \theta_g(t) + E[\max(A, t+\Delta)] \cdot (1 - \theta_g(t))$.

iii. Compute expected total risk
$C_g(t) = E[\varphi(T_{\text{conf}}) \mid t] + \lambda(1 - \theta_g(t))$.

(d) Find the optimal initial test week
$D_g^* = \arg\min_{t \in T_{\text{grid}}} C_g(t)$.

(e) Save the group policy $\pi_g = (D_g^*, \Delta)$.

Return $\mathrm{G}, \pi_g, p_g^{\text{succ}}, H_g$.

3. Model construction and optimization strategy

In the previous chapter, we explore how to scientifically group pregnant women based on their BMI and determine the appropriate initial NIPT testing time for each group to reduce potential risks and optimize retesting strategies. Building upon these analyses, this chapter will delve further into how to construct a data-driven model to refine and optimize the selection of NIPT testing times. We will explain in detail how to calculate the earliest threshold achievement time for each pregnant woman and, using survival analysis and machine learning techniques, provide personalized testing plans for different pregnant groups based on their BMI. The focus of this chapter is on establishing and solving a risk minimization model to achieve a more efficient and precise NIPT screening process [7].

3.1. Construction Method of the Earliest Threshold Achievement Time (A_i)

The raw data includes three types of information: "gestational week (including decimal)", "Y chromosome concentration", and "test success indicator (ok), determined by GC content". The steps to construct the earliest threshold achievement time (A_i) are as follows:

Threshold Judgment and Censoring Marking: Only the records with successful tests (ok=1) are used to determine if the threshold is met. For each pregnant woman, the gestational week in which the Y chromosome concentration first reaches or exceeds 4% is selected as the earliest threshold achievement time (A_i). If all successful test records of the pregnant woman show a Y chromosome concentration below 4%, A_i is considered "right-censored data," and a censoring indicator variable X_i is recorded (where $X_i = 1$ indicates achievement of the threshold, and $X_i = 0$ indicates censoring). This information will be used in subsequent survival model analysis. After calculation, the distribution of the earliest gestational weeks for male fetuses in pregnant women is shown in Figure 1.

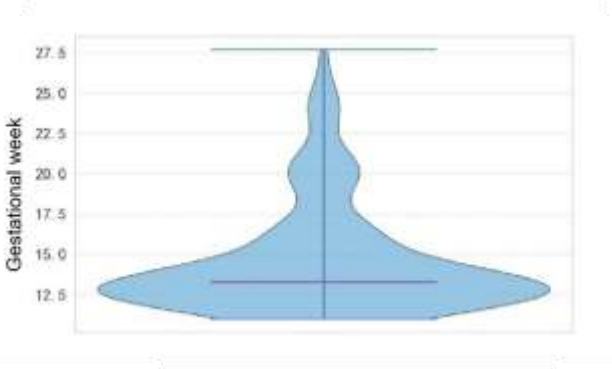

Fig 1. Distribution of estimated earliest gestational weeks for male fetuses among pregnant women

Noise Reduction and Monotonicity Correction: According to physiological principles, the Y chromosome concentration ($Y_i(t)$) of the pregnant woman should increase monotonically with gestational weeks. Therefore, the sequence of (gestational week t, Y chromosome concentration) for each pregnant woman needs to undergo isotropic monotonic regression or mild spline smoothing to ensure data smoothness. The first gestational week with a Y chromosome concentration $\geq 4\%$ is then selected from the corrected sequence as A_i, to avoid the occurrence of "false early achievement" (i.e., misjudging the threshold as reached too early due to data noise).

Handling of Failed Records: For test failure records (ok=0), although they do not participate in the "threshold achievement" judgment, their failure mechanisms (failure probability related to gestational week and BMI) need to be modeled separately. Ignoring failed records may underestimate the actual risk of premature testing, as premature testing may simultaneously face both the "threshold not met" and "sequencing failure" issues.

After completing the above calculations, we will plot the distribution of the earliest threshold achievement gestational weeks for male fetuses, showing the overall distribution characteristics of A_i.

3.2. Data-Driven BMI Grouping

Grouping Objective: The goal is to divide the continuous BMI variable into 3 to 5 intervals, ensuring that the groupings meet the following requirements: "few groups, feasible

implementation, stable group boundaries, and alignment with clinical needs." Each BMI group should correspond to a unified initial testing gestational week T_9 to ensure the practicality and effectiveness of the testing plan.

Grouping Method: A CART regression tree model is employed, using "earliest threshold achievement time A_i " as the response variable (the target to be predicted by the model) and "BMI" as the only predictor variable. In the regression tree, the number of leaf nodes K is limited (with a selectable range of $K \in 3,4,5,6$, and the optimal K is selected using five-fold cross-validation or the BIC criterion. On this basis, the model's default splitting criterion is replaced with the "risk reduction of the optimal initial detection gestational week o_g for each subset after splitting," so that the BMI group boundaries are directly linked to the goal of "minimizing expected risk." During the grouping process, priority is given to schemes with "fewer groups and stable boundaries" to avoid overfitting due to too many groups leading to small sample sizes, while ensuring the model's clinical applicability.

Grouping Results: Through decision tree classification analysis, the optimal grouping is determined to be 4 groups, with the Table 1.

Table 1. BMI division intervals for each group

Group	BMI
0	[26.81, 29.75)
1	[29.75, 32.68)
2	[32.68, 34.72)
3	[34.72, 43.71)

Model Error Analysis: The model errors for different K values are shown in Figure 2. The figure illustrates the average absolute errors (MAE, y-axis MAE = 2.38-2.44) corresponding to various K values from 4 to 9. This graph reflects the impact of changing the number of groups on model prediction accuracy, demonstrating that selecting the appropriate number of leaf nodes can optimize prediction accuracy and reduce overfitting.

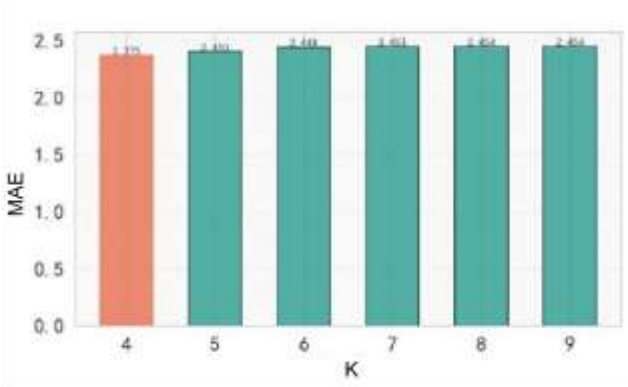

Fig 2. The impact of different K values on the MAE index of decision trees

3.3. Determination of the Optimal NIPT Testing Time for Each Group D_g

In this study, we design a strategy with one initial test and one retest, with a retest interval of X weeks. The following are the related foundational distributions and function definitions:

Basic Distribution and Function Definitions: Let $H_g(t)$ represent the conditional cumulative distribution function of the earliest threshold achievement time A_i in BMI group g (i.e., the probability of achieving the threshold before gestational week t). Its survival function is given by:

$$s_g(t) = 1 - H_g(t) \tag{3}$$

Expected Risk Calculation: The probability of confirming the test at time t is $p_0(t) = H_g(t)$, i.e., the probability of the fetus meeting the threshold at gestational week t . The expected risk of the confirmation time is given by the formula:

$$E_g(\varphi Q(t)) = \varphi(t) \cdot H_g(t) + E_g(\varphi \cdot \max(A_i, t+X)) \mid A_i > t \cdot s_g(t) \tag{4}$$

When the threshold is met at gestational week t , the risk is $\varphi(t)$; When the threshold is not met at gestational week t , the risk is the later time between A_i and the retest time $t+X$, weighted by the probability of not meeting the threshold $s_g(t)$.

Comprehensive Expected Risk: Considering the additional cost of sequencing failures φ , including the combined weighted value of time cost, economic cost, and maternal anxiety, the comprehensive expected risk function is:

$$C_g(t) = E_g(\varphi Q(t)) + \Phi \cdot s_g(t) \tag{5}$$

where Φ represents the cost of sequencing failure.

Solving for the Optimal Initial Testing Week o_g : Within the candidate gestational week range of "10–20 weeks" (with a step size of 0.5 weeks), we calculate the expected risk $C_g(t)$ for each group g at each candidate gestational week t . The gestational week o_g that minimizes the expected risk is chosen as the optimal initial testing time for that group.

Decision Trade-Off Logic. The earlier the testing time t is smaller, the early-stage risk $\varphi(t)$ is lower, but the probability of not meeting the threshold $s_g(t)$ is higher, leading to an increased probability of retesting and, consequently, higher costs for retesting and additional costs φ . The later the testing time t is larger: the probability of not meeting the threshold $s_g(t)$ is lower (higher probability of confirmation), but the risk $\varphi(t)$ may enter the "mid-stage (13-27 weeks, φ_2)" or "late-stage (≥ 28 weeks, φ_3)", where the potential risk increases significantly.

Retest Interval X and Additional Cost φ : The retest interval X and additional cost φ are critical factors that influence decision-making. Their values directly determine the decision tendency: whether to choose "early retesting to shorten waiting time" or "delay testing to pursue a successful first test." By considering these factors comprehensively, this study aims to find the most suitable initial testing time, optimizing the NIPT screening process.

The preliminary BMI grouping and recommended NIPT gestational weeks without error correction are shown in Table 2.

Table 2. Recommended results without error correction

Group	BMI	samples	Recommended first test for gestational age
0	[26.81, 29.75)	39	21.29
1	[29.75, 32.68)	110	12
2	[32.68, 34.72)	51	24.43
3	[34.72, 43.71)	36	24.72

4. Incorporating "Real Clinical Problems" into the Model Framework

4.1. Constructing the Error Model and Threshold Correction Under Measurement Errors

The observed Y chromosome concentration Y_{obs} is not the true value, but includes random observational noise η that follows a normal distribution with $N(0,\sigma^2)$. The relationship is as follows:

$$Y_{\text{obs}} = Y_{\text{true}} + \eta \tag{6}$$

In clinical practice, the 4% threshold may lead to false positives (incorrectly classifying a fetus as meeting the threshold) or false negatives (incorrectly classifying a fetus as not meeting the threshold) due to observational noise [8]. To enhance safety and eliminate these errors, we need to define an effective threshold:

$$4\%e = 4\% + z1-\alpha \cdot \sigma \quad (\text{where } \alpha = 0.05,, z_{0.95} = 1.645) \tag{7}$$

Then, using $4\%_e$, we recalculate the indicators A_i, $H_g(t)$, and O_g.

4.2. Estimating the Noise Variance σ

We propose two practical methods for estimating the variance of the observational noise σ:

Method 1: Use "short-interval repeated test data" from the same pregnant woman (with an interval between tests of ≤ 0.4 weeks, where the true Y chromosome concentration change can be neglected). The variance of the differences between two observed values can be used to indirectly estimate σ^2.

Method 2: Implement indirect control by "raising the threshold." Specifically, increase the threshold for classifying a result as "meeting the threshold." While this makes the threshold stricter and leads to a slightly later A_i, it stabilizes the "threshold achievement rate" and reduces fluctuations caused by noise.

4.3. Modeling the Mechanism of Sequencing Failure ok=0

Sequencing failure is not a random event, and its probability is usually related to factors such as gestational week, BMI, and others. This relationship needs to be quantified through a model:

Logistic Regression Model for Success Probability: We use binary logistic regression to model the sequencing success probability at a given gestational week t and BMI b. The formula is as follows:

$$\Pr(\text{success at } t|\text{BMI=b}) = \text{logit}^{-1}(d_0 + d_1 t + d_2 b + d_3 t \cdot b + \ldots) \tag{8}$$

From the coefficients, we can deduce the gestational week coefficient $d_1 > 0$: As the gestational week increases, the sequencing success probability increases. The BMI coefficient $d_2 < 0$: As BMI increases, the sequencing success probability decreases.

The interaction term coefficient d_3: Reflects the fact that in high-BMI groups, the effect of gestational week on the success probability weakens, i.e., as BMI increases, the positive impact of gestational week on success probability diminishes.

Comprehensive Probability of One-Time Confirmation: We define the probability of "one-time confirmation" as the probability that both "fetal Y chromosome concentration meets the threshold" and "sequencing is successful" are satisfied simultaneously. The comprehensive probability is defined as:

$$\theta_g(t) = \Pr(\text{success}|\text{t,g}) \times \text{H}_\text{g}(\text{t}|\text{success}) \tag{9}$$

Subsequently, we replace $H_g(t)$ with $\theta_g(t)$, and substitute it into the comprehensive expected risk function $C_g(t)$. Additionally, the cost of sequencing failure φ, such as the time cost of reappointments, economic costs, and maternal anxiety, is adjusted to reflect more clinically relevant values, optimizing the model's consideration of real-world costs:

$$C_g(t) = \varphi(t) \cdot \theta_g(t) + E[\varphi \cdot \max(A, t+\Delta) \mid A > t] \cdot (1-\theta_g(t)) + \varphi \cdot (1-\theta_g(t)) \tag{10}$$

4.4. Core Differences Between the Two Models

In the adjusted model, the "success confirmation rate for the current test" is no longer determined solely by "whether the fetus meets the threshold." Instead, it is determined by the product of the "threshold achievement probability" and the "sequencing success probability." If sequencing fails, not only is retesting required after the planned interval Δ weeks, but additional costs φ, including time for reappointments, economic costs, and maternal anxiety, must also be incurred.

Numerical Example: If, at 12 weeks, for a given group $H_g(12) = 0.8$ (80% of fetuses have met the threshold), $\Pr(\text{success}|12)=0.7$ (success probability is 70%). The overall success rate for the current test is:

$$\theta_g(12) = 0.8 \times 0.7 = 0.56 \tag{11}$$

This means that, at 12 weeks, the probability of "one-time confirmation" is 56%.

If, at 14 weeks and $H_g(14) = 0.95$ (95% of fetuses have met the threshold). $\Pr(\text{success}|14)=0.9$ (success probability is 90%). The overall success rate for the current test is:

$$\theta_g(14) = 0.95 \times 0.9 = 0.855 \tag{12}$$

At 14 weeks, 95% of fetuses have met the threshold, and the success rate is 90%, making the "one-time confirmation" probability increase to 85.5%. Compared to 12 weeks, the overall risk significantly decreases, and the retesting costs and waiting risks are also reduced.

4.5. Adjusted results

In this study, we recommended different initial NIPT testing weeks for different BMI groups and conducted a probability analysis with error correction. The results after error adjustment are shown in Table 3.

Table 3. Recommended results with error correction

Group	BMI	samples	Recommended first test for gestational age	Probability of compliance
0	[26.81,29.75)	37	20.0	0.865
1	[29.75,32.68)	94	12.0	0.170
2	[32.68,34.72)	44	20.0	0.932
3	[34.72,43.71)	24	19.6	0.500

The following are the specific recommendations and analysis results.

For Group 0 and Group 2, the recommended initial testing week is 20 weeks. The error-corrected probability of meeting the threshold on the first test is 0.865 for Group 0 and 0.932 for Group 2. This indicates that performing NIPT testing around the 20th week for these two groups of pregnant women can achieve a high "one-time confirmation rate" without frequent retesting, effectively controlling the additional costs and risks associated with retesting. This aligns with clinical requirements, ensuring testing accuracy while minimizing the frequency of retests.

For Group 3, the recommended initial testing week is 19.6 weeks, with an error-corrected probability of meeting the threshold on the first test being 0.500. Since both the "sequencing success probability" and the "Y chromosome threshold achievement probability" for high-BMI pregnant women rise more slowly, even during the 19.6-20 week testing window, only a clinically acceptable one-time confirmation rate can be achieved. Therefore, if clinical practice places more emphasis on "reducing retests" (i.e., a higher additional cost parameter φ, the candidate testing week can be adjusted to the 20-21 week range, and the comprehensive expected risk function can be recalculated to further optimize the testing time.

For Group 1, the recommended initial testing week is 12 weeks. However, the error-corrected probability of meeting the threshold on the first test is only 0.170, meaning most pregnant women will not achieve a one-time confirmation at 12 weeks. In this case, the unconfirmed probability, the additional cost of retesting, and the expected subsequent risk will increase significantly, contributing more to the comprehensive expected risk. Therefore, for this group, the focus should be on the rationality of the retest schedule, ensuring the optimization of retest timing and frequency to reduce unnecessary costs and risks.

5. Conclusion

This study develops a data-driven event-time optimization framework that integrates individualized feature analysis, risk segmentation, and cost-sensitive decision optimization. By grouping samples according to feature-based intervals and recommending optimal initial decision points for each group, the proposed model enables adaptive and efficient scheduling under uncertainty. For cases where the threshold condition is not satisfied during the first evaluation, a secondary decision strategy is introduced to balance accuracy, cost, and risk, ensuring computational efficiency and robustness. The model further quantifies how individual covariates influence event-time distribution and incorporates an error-correction mechanism to enhance the reliability of threshold-based predictions.

Experimental analysis demonstrates that the optimized decision timing derived from the proposed model corresponds to the highest probability of one-step confirmation across most feature groups, minimizing unnecessary repetitions and associated costs. Beyond its biomedical data origin, the framework provides a generalizable modeling paradigm for event-time prediction, dynamic decision scheduling, and big-data-driven optimization across diverse domains such as predictive maintenance, intelligent monitoring, and reliability modeling.

References

[1] Rose, N. C., Barrie, E. S., Malinowski, J., Jenkins, G. P., McClain, M. R., LaGrave, D., ... & ACMG Professional Practice and Guidelines Committee. (2022). Systematic evidence-based review: the application of noninvasive prenatal screening using cell-free DNA in general-risk pregnancies. Genetics in Medicine, 24(7), 1379-1391.

[2] Dungan, J. S., Klugman, S., Darilek, S., Malinowski, J., Akkari, Y. M., Monaghan, K. G., ... & ACMG Board of Directors. (2023). Noninvasive prenatal screening (NIPS) for fetal chromosome abnormalities in a general-risk population: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). Genetics in Medicine, 25(2), 100336.

[3] Cernat, A., De Freitas, C., Majid, U., Trivedi, F., Higgins, C., & Vanstone, M. (2019). Facilitating informed choice about non-invasive prenatal testing (NIPT): a systematic review and qualitative meta-synthesis of women's experiences. BMC pregnancy and childbirth, 19(1), 27.

[4] Oliveri, S., Ongaro, G., Cutica, I., Menicucci, G., Belperio, D., Spinella, F., & Pravettoni, G. (2023). Decision-making process about prenatal genetic screening: how deeply do moms-to-be want to know from Non-Invasive Prenatal Testing?. BMC pregnancy and childbirth, 23(1), 38.

[5] Zheng, J., Lu, H., Li, M., Guan, Y., Yang, F., Xu, M., ... & Zhou, Y. (2020). The clinical utility of non-invasive prenatal testing for pregnant women with different diagnostic indications. Frontiers in Genetics, 11, 624.

[6] Kater-Kuipers, A., De Beaufort, I. D., Galjaard, R. J. H., & Bunnik, E. M. (2020). Rethinking counselling in prenatal screening: an ethical analysis of informed consent in the context of non-invasive prenatal testing (NIPT). Bioethics, 34(7), 671-678.

[7] Allyse, M., Sayres, L. C., Goodspeed, T., Michie, M., & Cho, M. K. (2015). "Don't want no risk and don't want no problems": public understandings of the risks and benefits of noninvasive prenatal testing in the United States. AJOB empirical bioethics, 6(1), 5-20.

[8] Schöne-Seifert, B., & Junker, C. (2021). Making use of non-invasive prenatal testing (NIPT): rethinking issues of routinization and pressure. Journal of Perinatal Medicine, 49(8), 959-964.