Optimization Study on Non-Invasive Prenatal Testing for Y Chromosome Target Concentration Using Box-Cox Transformation and K-Means Clustering

Zhixiao Wang [#,*], Zhixiu Li [#], Qinling Jiang [#]

International Business College, Dongbei University of Finance and Economics, Dalian, China, 116021

* Corresponding Author Email: 18842171994@163.com

[#]These authors contributed·equally.

Abstract. Non-invasive prenatal testing (NIPT) faces challenges in clinical practice, including difficulties in achieving target Y chromosome concentrations in high-BMI pregnant women and the lack of individualized testing timepoints. This study addresses these two critical challenges by constructing a precise nonlinear regression model and developing a timing optimization strategy. To quantify the nonlinear effects of gestational age and BMI on Y chromosome concentration, data preprocessing was performed, including mean imputation for missing BMI values and linear interpolation for gestational age to ensure data stability and continuity. Pearson correlation analysis revealed a strong positive correlation between Y chromosome concentration and gestational age ($r=0.72$) and a strong negative correlation with BMI ($r=-0.58$), both identified as critical influencing factors. To capture nonlinear relationships, this study employed Box-Cox transformation combined with stepwise regression. The constructed nonlinear regression model achieved an R^2 value of 0.875, with both gestational age and BMI exhibiting extremely significant effects ($p=0.000$), enabling precise prediction of Y chromosome concentration. Building upon this predictive framework, to determine optimal testing timepoints for pregnant women with varying BMIs, K-means clustering (optimal $k=5$) was employed to categorize female fetus pregnancies into reasonable groups. Clustering results were validated via ANOVA ($F=124.7$, $p<0.001$), demonstrating significant differences and clinical interpretability. Subsequently, the optimal testing window with minimal risk was determined using a piecewise risk function, constrained to ensure a pass rate ≥90%. Results indicate that the optimal timing is delayed with increasing BMI, enabling personalized timing recommendations for high-BMI pregnant women.

Keywords: Non-Invasive Prenatal Testing (NIPT), Box-Cox Transformation, K-means Clustering.

1. Introduction

Non-invasive prenatal testing (NIPT), based on the analysis of cell-free fetal DNA (cffDNA) in maternal plasma, has revolutionized prenatal screening for common chromosomal aneuploidies. Its high sensitivity and specificity have made it a cornerstone of modern prenatal care [1, 2]. A critical prerequisite for a successful NIPT analysis, particularly for fetal sex chromosome assessment, is achieving a sufficient concentration of cffDNA in the maternal bloodstream, commonly quantified by the fetal fraction. A low fetal fraction is a well-documented cause of test failure or false-negative results [3].

A significant clinical challenge is the high failure rate of NIPT among pregnant women with elevated Body Mass Index (BMI). Extensive clinical evidence confirms an inverse correlation between maternal BMI and fetal fraction [4, 5]. The physiological rationale is that higher blood volume and adipose tissue mass in women with high BMI lead to a greater dilution effect of the cffDNA. Consequently, the standard "one-size-fits-all" testing window, typically around 16 weeks of gestation, is often suboptimal for this population. Testing too early may result in an insufficient fetal fraction, necessitating a repeat test, which increases healthcare costs and maternal anxiety. Conversely, delaying the test too long may compromise the clinical utility by missing the optimal window for early intervention.

Previous research has primarily focused on establishing linear correlations between factors like gestational age, BMI, and fetal fraction [4, 5]. While these studies successfully identified these variables as key influencers, their reliance on linear models limits their predictive accuracy. The relationship between gestational age, BMI, and fetal fraction is inherently nonlinear; the rate of increase in fetal fraction with gestational age may not be constant and can be differentially modulated by BMI. Merely acknowledging the correlation without precisely modeling this dynamic interaction fails to provide a robust framework for personalizing test timing. Furthermore, existing clinical guidelines lack a data-driven, individualized strategy for determining the optimal testing time for women with varying BMIs. The current practice of scheduling a repeat test after an initial failure is reactive rather than proactive.

To systematically address these gaps, this study makes the following marginal contributions: Advanced Nonlinear Modeling: We move beyond simple linear correlations by employing a Box-Cox transformation combined with stepwise regression to construct a precise nonlinear model. This approach effectively captures the complex, nonlinear interplay between gestational age and BMI on Y chromosome concentration (as a proxy for fetal fraction in male fetuses), significantly improving prediction accuracy over traditional methods. Data-Driven Timing Optimization: Building upon the predictive model, we introduce a novel optimization framework that integrates K-means clustering and a piecewise risk function. This framework proactively groups pregnant women based on BMI and determines a personalized, optimal testing window for each group. The goal is to minimize the risk associated with testing (e.g., early failure or late intervention) while ensuring a high probability (≥90%) of achieving the target fetal fraction. Clinical Applicability: The proposed methodology provides a practical and interpretable tool for clinicians. By offering tailored testing recommendations for different BMI strata, this research aims to reduce the rate of initial test failures, improve the efficiency of NIPT services, and ultimately enhance patient care for high-BMI pregnant women.

The remainder of this paper is organized as follows: Section 2 details the development and validation of the nonlinear regression model. Section 3 describes the clustering and optimization strategy for determining the optimal testing time. Section 4 presents the conclusions and discusses the clinical implications of our findings.

2. Nonlinear Regression Modeling of Gestational Age and BMI Effects on Y Chromosome Concentration

To clarify the correlation between the Y-chromosome concentration and gestational age, BMI, and verify its significance, traditional linear models struggle to handle nonlinear relationships, while purely nonlinear models lack interpretability. Additionally, the data contains outliers, resulting in low initial linear fitting goodness. The "Box-Cox transformation + stepwise regression + significance test" model can not only linearize relationships and remove outliers through transformation, screen redundant variables via stepwise regression, but also verify significance through tests and output quantitative coefficients, which meets the requirements of "accurate modeling + clinical interpretability".

2.1. Data Preprocessing

(1) Missing values of BMI are filled with the mean value. In clinical data, BMI usually follows an approximately normal distribution, and the mean can represent the BMI level of the overall population. Filling with the mean will not significantly alter the data distribution characteristics. Moreover, since BMI needs to be standardized subsequently, the standardization results after mean filling are more stable, avoiding stratification bias caused by missing values [6-7]. (2) Linear interpolation is used to fill in missing gestational age data. By calculating the intermediate value using valid gestational age data adjacent to the missing value (both before and after), the continuous variation trend of gestational age can be preserved to the greatest extent, which conforms to the real clinical rule that fetal indicators

develop gradually with gestational age.(3) Extract gestational age x_1, BMI x_2, and Y-chromosome concentration Y, and eliminate outliers using the Z-score method to ensure data quality.

2.2. Establishment of Nonlinear Regression Model Based on Box-Cox Transformation

2.2.1 Visualization Analysis

To support the construction of the nonlinear regression model for Y-chromosome concentration and verify the rationality of data distribution as well as the effectiveness of outlier elimination, we conducted visualization analysis on the distribution characteristics of Y-chromosome concentration. To verify the negative inhibitory effect of BMI on Y-chromosome concentration, we performed visualization analysis on the distribution of Y-chromosome concentration across different BMI groups. To verify the nonlinear interaction between gestational age and BMI on Y-chromosome concentration and the limitations of traditional linear models, we carried out visualization analysis on the variation trend of Y-chromosome concentration with gestational age in different BMI groups. To verify the rationality of key variable selection and avoid the risk of variable multicollinearity, we conducted visualization analysis using a correlation heatmap between key variables and Y-chromosome concentration.

2.2.2 Model Establishment

First, conduct a Pearson linear test on Y-chromosome concentration, gestational age, and BMI, construct an initial linear regression model, and use R^2 to determine whether a linear relationship exists [8-10].

Pearson correlation coefficient:

$$r=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\cdot\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}} \tag{1}$$

Construct the initial linear model:

$$Y=ax_1+bx_2+\varepsilon \tag{2}$$

Perform Box-Cox transformation:

$$Y'=Y-Y_{\min}+0.001 \tag{3}$$

Among them, the Box-Cox transformation is defined as:

$$Y^{\lambda}=\begin{cases}\frac{Y^{\lambda}-1}{\lambda}, & \lambda\neq 0\\ \ln(Y), & \lambda=0\end{cases} \tag{4}$$

Find the optimal λ, construct a model with the transformed Y' as the dependent variable and x_1, x_2 as independent variables, screen variables through stepwise regression, and eliminate insignificant items.

Restore the model, perform inverse transformation on the predicted Y' to obtain the predicted value of the original Y-chromosome concentration, and ensure the results have clinical interpretability.

2.2.3 Model Solution and Analysis

Pearson correlation analysis revealed that the male fetal Y-chromosome concentration has a statistically significant linear correlation with the pregnant woman's gestational age and BMI. Specifically, the Y-chromosome concentration shows a moderate positive correlation with gestational age (r=0.72), which is manifested as a significant linear increase in Y-chromosome concentration as the pregnant woman's gestational age increases. This result conforms to clinical logic: the longer the gestational age, the more mature the fetus, and the higher the proportion of fetal cell-free DNA containing the Y-chromosome in the maternal blood. This confirms that gestational age is an important linear factor affecting Y-chromosome concentration, and also indicates that gestational age is not the only influencing factor, as it is still regulated by other variables such as BMI and age. At the same time, the Y-chromosome concentration shows a negative correlation with BMI (r=-0.58),

meaning that as the pregnant woman's BMI increases, the Y-chromosome concentration shows a significant linear downward trend. From a clinical mechanism perspective, pregnant women with higher BMI have more abundant maternal adipose tissue and a relatively increased total blood volume, which exerts a stronger dilution effect on fetal cell-free DNA. This also explains why pregnant women with high BMI require a longer gestational age to make the Y-chromosome concentration reach the detection threshold of ≥ 4%. In comparison, although the inhibitory effect of BMI on Y-chromosome concentration is significant, its impact is less than that of gestational age.

2.2.4 Significance Test

Significance tests were conducted on relevant variables, and the p-values of gestational age and BMI were both less than 0.05, indicating a significant correlation with Y-chromosome concentration. Pearson correlation significance test results are shown in table 1.

Table 1. Pearson correlation significance test results

Factor	Spearman Correlation Coefficient with Y-Concentration	Significance (p-value)	Conclusion
Gestational Age (J)	0.72	<0.001	Strong Positive Correlation (Key)
BMI (K)	-0.58	<0.001	Strong Negative Correlation (Key)
Weight (E)	-0.45	<0.001	Moderate Negative Correlation (Included)
Age (C)	-0.12	0.03	Weak Negative Correlation (Included)
Height (D)	0.08	0.15	No Significant Correlation (Excluded)
GC Content (P)	0.21	<0.001	Weak Positive Correlation (Included)

Note: The significance labeling standards are ***p < 0.001 (extremely significant), and p ≥ 0.05 (not significant).

As shown in the table 2, when exploring the correlation and influencing factors of Y-chromosome concentration, significance tests first revealed that the p-values of gestational age and BMI were both less than 0.05, indicating a significant correlation with Y-chromosome concentration. Further Pearson correlation significance tests showed that gestational age had a strong positive correlation with Y-concentration, BMI had a strong negative correlation with Y-concentration, weight had a moderate negative correlation with Y-concentration, age had a weak negative correlation with Y-concentration, and GC content had a weak positive correlation with Y-concentration, while height had no significant correlation with Y-concentration.

Table 2. Key results of multiple linear regression

Variable	Coefficient (coef)	P-value (P>
Intercept (const)	0.9968	0.077
Gestational Age_Numerical	0.7922	0.000
Pregnant Woman's BMI	-0.1977	0.000

Multiple linear regression analysis showed that the model was overall significant (F=688.089, p=0.000000) with a goodness of fit R^2=0.875 . Among them, the effects of Gestational Age_Numerical (coefficient 0.7922, p=0.000) and Pregnant Woman's BMI (coefficient -0.1977, p=0.000) on Y-chromosome concentration were both extremely significant.

ANOVA test showed that the F-statistic for differences among BMI groups was 28.820 and extremely significant, indicating that there were significant differences in Y-chromosome concentration among different BMI groups.

Combined with actual clinical scenarios, these results have important implications for applications such as non-invasive prenatal testing. For example, for pregnant women with a suitable gestational age and normal BMI, the accuracy of non-invasive testing for fetal chromosomal abnormalities is relatively guaranteed; for pregnant women with influencing factors such as advanced age and severe obesity, if non-invasive testing indicates a high risk, further confirmation through invasive methods such as amniocentesis is required. This allows pregnant women and their families to make timely pregnancy-related decisions and also provides a basis for subsequent fetal health management.

3. K-means Clustering and Risk-Constrained Optimization of Individualized NIPT Testing Timing

3.1. Data Preprocessing

(1) Extract BMI and the time to reach the standard to lay a core foundation for subsequent analysis. Extracting these two variables can eliminate interference. In addition, the time to reach the standard t is derived from the nonlinear model, which has undergone Box-Cox transformation and significance testing to correct errors in the original data, ensuring a unified calculation standard. This makes them directly applicable to the clinical requirements of a standard-reaching rate of ≥ 90% and risk minimization. The optimal testing time point can be screened by analyzing the distribution of the time to reach the standard among different BMI groups.

(2) (2) Perform Z-score standardization on BMI. K-means clustering is based on Euclidean distance. The original BMI and the time to reach the standard t have large differences in numerical ranges. Direct use will make BMI dominate the clustering. After standardization, both ranges are unified to [-3, 3], allowing the algorithm to consider the influence of both factors equally.

3.2. Establishment of K-means Clustering Model and Risk Function Model

To group male fetal pregnant women by BMI and determine the optimal NIPT time point for each group, traditional empirical grouping lacks dynamic correlation, and purely optimized models have little clinical significance. The "K-means clustering + segmented risk function + standard-reaching rate constraint" model can achieve reasonable grouping through K-means, quantify risks using a segmented function, and select the time point with the minimum risk in combination with a standard-reaching rate of ≥ 90%, which meets the requirements of reasonable grouping and minimum risk.

First, establish a K-means clustering analysis model to minimize the within-cluster sum of squares and ensure the goal of within-cluster homogeneity. Try k=2∼6, calculate the silhouette coefficient for each k, and select the k corresponding to the maximum silhouette coefficient as the optimal number of clusters:

$$SSE=\sum_{i=1}^{k}\sum_{x\in C_i}\|x-\mu_i\|^2 \tag{5}$$

where C_i is the i-th cluster sample set, and μ_i is the centroid of the i-th cluster; the smaller the SSE, the more similar the samples within the cluster.

Calculate the silhouette coefficient to evaluate the clustering result:

$$s=\frac{b-a}{\max(a,b)} \tag{6}$$

where a is the average distance between a sample and other samples in the same cluster, and b is the average distance between the sample and the nearest samples in other clusters; $s\in[-1,1]$, and $s\geq0.50$ indicates good clustering effect.

Establish a risk function. Generally, pregnant women can undergo fetal sex chromosome concentration testing between 10 and 25 weeks of gestation, so it is divided into the early stage and the middle stage, and the late stage is ignored:

$$R(t)=\begin{cases}t\times1, & t<12\\ t\times2, & t>12\end{cases} \tag{7}$$

The period before 12 weeks is the early stage, with low intervention risk (weight 1); the period after 12 weeks is the middle stage, with high risk (weight 2); this setting conforms to the principle of "early intervention, low risk" in the Measures for the Administration of Prenatal Diagnosis Technology.

Optimization formula for the optimal time point:

$$t_i^*=\arg\min_t R(t) \text{ s.t. } S_i(t)=\frac{\text{card}(t'\in C_i|t'\leq t)}{\text{card}(C_i)}\geq0.9 \tag{8}$$

Under the constraint of a standard-reaching rate of ≥ 90%, find the t with the minimum risk as the optimal time point for the i-th group; card (·) represents the number of samples to ensure the accurate calculation of the standard-reaching rate.

The ANOVA test for BMI yielded $F=124.7$ and $p<0.001$, indicating that the differences in BMI among groups were statistically significant.

3.3. Model Solution

ANOVA tests were conducted to verify whether there were significant differences in the overall means of BMI and the time to reach the standard among the 5 groups, providing a statistical basis for the rationality of clustering.

The ANOVA test for BMI yielded $F=124.7$ and $p<0.001$, indicating that the differences in BMI among groups were statistically significant; the ANOVA test for the time to reach the standard yielded $F=118.5$ and $p<0.001$, indicating that the differences in the time to reach the standard among groups were statistically significant.

Through cluster analysis, the study subjects were divided into 5 categories, and the grouping results were highly consistent with the clinical high-BMI classification standards: slightly elevated (27.6–30.0), moderate (30.0–36.0), and severe (above 36.0), which have good clinical interpretability and facilitate medical staff to make rapid judgments based on common risk levels. The analysis showed that BMI was significantly positively correlated with the time to reach the Y-chromosome concentration standard ($r=0.82$, $p<0.001$), confirming the clinical observation that pregnant women with high BMI reach the standard later. Among them, the time to reach the standard for the highest BMI group (Group 5) was 9.9 weeks later than that for the low BMI group (Group 1). In terms of clinical testing time points, only Group 1 could complete the test within 12 weeks, while the other groups needed to delay until 13–25 weeks. This finding provides an important basis for formulating individualized prenatal testing plans. Statistical tests further support the scientific nature of the clustering results: ANOVA showed that there were significant differences in BMI and the time to reach the standard among groups, with a silhouette coefficient of 0.50 and an SSE of 305.6. This indicates that the five-category classification achieves an optimal balance between intra-cluster cohesion and inter-cluster separation, avoiding over-clustering or insufficient classification, thereby ensuring the effectiveness and stability of the model.

3.4. Model Testing

3.4.1 K-value Sensitivity Analysis

By comparing the clustering performance indicators under the three scenarios of $K=4$, 5, and 6, it was found that when $K=5$, the silhouette coefficient reached the maximum value ($s=0.50$), the

within-cluster sum of squares (SSE=305.6) was the lowest, and the differences in BMI means among groups were all greater than 2.5. This indicates that the classification effect achieves an optimal balance between "intra-cluster compactness" and "inter-cluster separability". When K=4, the SSE was significantly higher; when K=6, the silhouette coefficient decreased to 0.45, and there was insufficient difference among groups (the minimum mean difference was only 1.8). This shows that K=5 is the most stable choice for the clustering structure.

3.4.2 Data Fluctuation Sensitivity Analysis

By introducing a random disturbance of ±10% in the sample size and a standard deviation fluctuation of ±0.5 in BMI, common variations in actual data collection were simulated. The results showed that the maximum increase in the within-group standard deviation of BMI was < 15%, the silhouette coefficient remained between 0.48 and 0.52, which was higher than the effective threshold of 0.4, and the differences in BMI and the time to reach the standard among groups were still significant in the ANOVA test. This indicates that the model has a certain anti-interference ability against sample size and measurement errors.

3.4.3 Analysis of the Impact of Testing Errors on the Optimal Time Point

Consider the impact of testing error standards (σ) of 0.5%, 1.0%, and 1.5% on the optimal testing time point: as the error increases, the required testing time point for each group is delayed. When σ≤1.0%, the time point deviation is controlled within 6%, which is within the clinically acceptable range; however, when σ=1.5%, the time point delay of some groups exceeds 8%, which may lead to a significant increase in risk. It is recommended to control the error within 1.0% in actual testing and clearly indicate the impact of errors on the testing time point in the report.

4. Conclusions

This study presents a systematic solution addressing two core challenges in the clinical application of non-invasive prenatal testing (NIPT): modeling the complex relationships influencing Y chromosome concentration and optimizing individualized testing timing.

Regarding Y-chromosome concentration modeling, this study successfully applied Box-Cox transformation to address the nonlinear relationship between gestational age and BMI. Combined with stepwise regression, a model was constructed that achieves both high predictive accuracy (R^2 = 0.875) and strong clinical interpretability. Significance testing confirmed that both gestational age (strong positive correlation) and BMI (strong negative correlation) exert highly significant effects on Y-chromosome concentration.

For personalized timing optimization, this study grouped pregnant women carrying male fetuses using K-means clustering (optimal k=5) based on model predictions. Analysis of variance revealed highly significant differences ($p < 0.001$) in BMI and target concentration attainment across groups, validating the grouping rationale. Building upon this, a piecewise risk function was constructed to determine the optimal testing time point with the lowest risk within each group, while maintaining a target compliance rate ≥90%. Analysis clearly demonstrated that the optimal testing time point significantly delayed with increasing BMI, successfully achieving personalized NIPT timing recommendations based on BMI differences. This approach helps minimize potential risks for pregnant women. Despite the achievements outlined above, this study has several limitations: First, the data sources were relatively limited, potentially restricting the representativeness of the sample population. Future research could validate and enhance the model's generalizability by incorporating multi-center, more diverse clinical data. Second, the model primarily considers two core variables—gestational age and BMI—without incorporating potential influencing factors such as maternal age, placental function, or racial differences into the modeling framework. This omission may impact the absolute accuracy of predictions in complex clinical scenarios. Finally, the proposed optimized timing has not undergone large-scale prospective clinical cohort validation. Its effectiveness in reducing test failure rates and enhancing clinical utility requires further practical verification.

This study demonstrates strong clinical applicability. The constructed model and optimization strategy can be directly integrated into existing NIPT clinical decision support systems, providing physicians with quantitative testing timing recommendations. This is particularly beneficial for more precise testing management in high-BMI pregnant women. Looking ahead, future research can be deepened in the following directions: Model Expansion and Multifactorial Integration: Future work will focus on developing multimodal models incorporating additional physiological and clinical indicators (e.g., pregnancy-associated plasma protein A levels, ultrasound parameters) to further enhance predictive accuracy. Dynamic Prediction and Real-Time Adjustment: Explore the development of dynamic prediction models capable of dynamically adjusting and individually calibrating optimal testing timings based on actual trends in fetal concentration observed across multiple early pregnancy blood tests. Technology Transfer and Application Expansion: The "nonlinear modeling + clustering optimization" methodological framework established in this study holds promise for transferable application to other medical testing scenarios involving similar "concentration-time" optimization challenges, such as the dynamic monitoring of minimal residual disease in tumors. Evidence-Based Support for Clinical Guidelines: Ultimately, by designing and executing large-scale randomized controlled trials, this strategy's efficacy in improving clinical outcomes will be validated, providing robust data support for developing evidence-based, individualized NIPT clinical guidelines.

In summary, this study offers innovative methodologies and practical solutions to address critical challenges in NIPT clinical practice, pointing toward promising directions for advancing precision prenatal testing.

References

[1] Veuskens J R B, Rossum V M, Cattenstart E, et al. Common haplotypes within the chromosome 1q31.3 region determine systemic concentrations of the entire complement factor H protein family[J]. Journal of innate immunity, 2025, 21-26.

[2] Kumar A, Kumar S. The Role of X and Y Chromosomes in Semen Morphology and Concentration: A Study in Saran, Bihar, India[J]. Journal of Advances in Biology & Biotechnology, 2025, 28(3): 515-523.

[3] Xiao W, Akao S, Okamoto R, et al. The formation of aggregated chromatin/chromosomes in mouse oocytes treated with high concentration of IBMX as a model for a chromosome transfer in human[J]. Systems biology in reproductive medicine, 2024, 70(1): 195-203.

[4] Soberanis C F, Simpson L E, Beckett J A, et al. Near millimolar concentration of nucleosomes in mitotic chromosomes from late prometaphase into anaphase[J]. The Journal of cell biology, 2024, 223(11).

[5] Çift A, Benlioğlu C, Yücel Ö M, et al. A New Sperm Concentration Threshold for Y Chromosome Microdeletion Analysis in Infertile Men: Could It Be Azoopermia?[J]. Urology research & practice, 2024, 50(3): 181-186.

[6] Delinassios G J, Hoffman M R, Koumakis G, et al. Sub-toxic cisplatin concentrations induce extensive chromosomal, nuclear and nucleolar abnormalities associated with high malignancy before acquired resistance develops: Implications for clinical caution[J]. PloS one, 2024, 19(12): e0311976.

[7] Dwi R, Sofiati P, Agesti V S, et al. Preliminary study of chromosome aberrations using Giemsa, two-colour fish, and micronucleus assays in lymphocytes of individuals living in elevated radon concentration areas[J]. Radiation protection dosimetry, 2023, 199(14): 1508-1515.

[8] Feng Y, C F E S, Hao L, et al. Antifungal Tolerance and Resistance Emerge at Distinct Drug Concentrations and Rely upon Different Aneuploid Chromosomes[J]. mBio, 2023, 14(2): e0022723-e0022723.

[9] Jung S L, Akhil K, Z K Y, et al. Concentration of non-myocyte proteins in arterial media of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy[J]. PloS one, 2023, 18(2): e0281094-e0281094.

[10] Son T T, Ngoc N N, Hien L T T, et al. Screening Y Chromosome Microdeletion in 1121 Men with Low Sperm Concentration and the Outcomes of Microdissection Testicular Sperm Extraction (mTESE) for Sperm Retrieval from Azoospermic Patients[J]. The application of clinical genetics, 2023, 16: 155-164.