

Received: 2021.11.22

Accepted: 2022.02.04

Available online: 2022.02.21

Published: 2022.04.13

External Quality Assessment of Maternal Serum Levels of Alpha-Fetoprotein, Free Beta-Human Chorionic Gonadotropin, and Unconjugated Estriol in Detecting Down Syndrome and Neural Tube Defects in the Second Trimester of 87 Maternal Serum Samples, Based on 105-139 Days

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

AEG 1,2 **Yiming Chen*** 
BE 2 **Yijie Chen*** 
CD 3 **Yezhen Shi**
BE 2 **Wenwen Ning**
B 1 **Xiaoying Wang**
BC 1 **Liyao Li**
F 2 **Huimin Zhang**

1 Department of Prenatal Diagnosis and Screening Center, Hangzhou Women's Hospital (Hangzhou Maternity and Child Health Care Hospital), Hangzhou, Zhejiang, PR China
2 The Fourth School of Clinical Medicine, Zhejiang Chinese Medical University, Hangzhou, Zhejiang, PR China
3 Department of Data Analysis, Zhejiang Bosheng Biotechnology Co Ltd., Hangzhou, Zhejiang, PR China

* Yiming Chen and Yijie Chen contributed equally to this work

Corresponding Author: Yiming Chen, e-mail: cxy40344@163.com

Financial support: This study was supported by the Zhejiang Public Welfare Technology Research, Program/Social Development (grant no. LGF19H040006) and the Hangzhou Health, Science, and Technology Plan Project (no. 2017A055)

Conflict of interest: None declared

Background: We aimed to insure the accuracy and reproducibility of alpha-fetoprotein (AFP), free beta-human chorionic gonadotropin (free β -hCG), and unconjugated estriol (uE3) concentrations for the screening for trisomy 21 (T21) and neural tube defects (NTD) in the second trimester. We conducted an external quality assessment of 6 laboratories, using maternal serum specimens.

Material/Methods: Serum specimens collected from 87 women of singleton pregnancies (4 with T21, 5 with NTD, and 78 with normal fetuses) were divided into 6 equivalent-volume fractions and transported to 6 laboratories (A, B, C, D, E, and F). All laboratories used the time-resolved fluorescence analyzer and supporting reagents to measure concentrations of AFP, free β -hCG, and uE3. The screening efficacies of T21 and NTD were compared with the certified or accredited status of the participants' quality systems.

Results: Concentrations of AFP measured by laboratory F were low compared with those determined by the other 5 laboratories, and the differences were significant ($P < 0.01$). There was no statistically significant difference in the free β -hCG and uE3 concentrations measured by the 6 laboratories ($P > 0.05$). The correlation coefficients for the 3 multiples of the median values were all > 0.900 . The McNemar paired chi-squared test showed the differences in the positivity and detection rates were not statistically significant ($P = 1.000$).





Conclusions: AFP, free β -hCG, and uE3 values measured by the other 5 laboratories were comparable with those of laboratory A, with good linear correlation. When used in the maternal prenatal screening of T21 and NTD, the test results met the clinical requirements.

Keywords: **alpha-Fetoproteins • Chorionic Gonadotropin, beta Subunit, Human • Down Syndrome • Estriol • Neural Tube Defects**

Abbreviations: **T21** – trisomy 21; **NTD** – neural tube defects; **AFP** – alpha-fetoprotein; **free β -hCG** – free beta-human chorionic gonadotropin; **uE3** – unconjugated estriol; **PPV** – positive predictive value; **NPV** – negative predictive value; **FPR** – false positive rate; **FNR** – false negative rate; **PR** – positive rate; **DR** – detection rate

Full-text PDF: <https://www.medscimonit.com/abstract/index/idArt/935573>



 3216  10  5  38

Background

Alpha-fetoprotein (AFP), free beta-human chorionic gonadotropin (free β -hCG), and unconjugated estriol (uE3) are used in the maternal prenatal screening of trisomy 21 (T21, also known as Down syndrome) and neural tube defects (NTD) in the second trimester of pregnancy [1-3]. A meta-analysis showed that double and triple tests involving AFP, uE3, total hCG, and free β -hCG were significantly better than single marker tests. When the false positive rate (FPR) is 5%, 60% to 70% of Down syndrome could be detected [4]. Fetuses with T21 are relatively immature compared with fetuses without T21, so the level of hCG will be higher than expected, while the levels of AFP and uE3 will be lower than expected [5]. Studies have shown that the best time to measure maternal serum AFP for detecting NTD is at 16 to 18 weeks of pregnancy [6]. Individuals with Down syndrome, which is the most common chromosomal abnormality, have a significantly higher risk of congenital cardiac disease, Alzheimer disease, and autoimmune disease [7]. It has been demonstrated that the prevention, screening, and diagnosis of NTD can significantly improve the outcome of mothers and infants. However, mothers and infants can be harmed if the patient refuses routine ultrasound scans or delays critical care milestones for T21 and NTD screening, adversely affecting the patient, infant, patient's family, and healthcare system [8].

Prenatal screening findings can affect the decision of whether to continue a pregnancy [9]. As such, it is the physician's responsibility to provide sufficient information so that an informed decision can be made by the patient and the patient's family. Thus, prenatal screening has an essential role in identifying common genetic disorders [10-12].

Clinical research laboratories require high-quality and high-quantity specimens to generate reliable test results [13]. Quality assurance refers to the steps taken to ensure the reliability and improve the accuracy of the test results according to national and international laws as well as other standards and guidelines. These steps include internal quality control procedures and external quality assessment (EQA) methods that involve the pre-analytic process, target analyte type, specimen type, platform type, test components, and positive and negative controls [14,15].

To ensure the accuracy and the reproducibility of the test results, EQA should be implemented from the pre-analytic stage to the analytic stage [16]. To examine the accuracy and the reproducibility of T21 and NTD screening results across 6 laboratories using the same instrument and the same analysis, we measured the concentrations of AFP, free β -hCG, and uE3. The results and risk values were compared and studied to provide a foundation for a standardized management system for maternal prenatal screening in Hangzhou.

Material and Methods

Screening Laboratory Conditions

The 6 laboratories were designated as A, B, C, D, E, and F. Laboratory A had the largest sample size and was one of the earliest laboratories in the region, which was identified by the regional pre-production screening Quality Control Management Center. The laboratory quality assessment results of the national inspection center are qualified every year. Thus, laboratory A was the reference laboratory with which all other laboratories were compared. The second trimester maternal serum levels of AFP, free β -hCG, and uE3 were measured.

Reagents and Instruments

The 1235 automatic time-resolved fluorescence analyzer (PerkinElmer, Shelton, CT, USA) and Lifecycle 4.0 software were used in all laboratories. An AFP/free β -hCG double-labeling kit and uE3 reagent supporting Wallac software (Wallac Oy, Turku, Finland) were used in conjunction with quality control samples (batch no. UF2181101, Hangzhou Biozone Technology Co., Ltd., Hangzhou, China).

Inclusion and Exclusion Criteria

There were 87 randomly selected cases of prenatal screening in laboratory A. The inclusion criteria were single pregnancy and voluntarily serological screening in the second trimester. There were specimens from pregnant women from gestational age of 119 (105-139) days, maternal age of 28.13 (19.52-39.12) years, and maternal weight of 57 (44.52-74.20) kg. Among them, there were 4 women with serum samples with T21 and 5 with NTD and 78 were normal. The exclusion criteria were as follows: double or multiple pregnancies, smoking, diabetes, history of chromosomal abnormalities and congenital malformation, in vitro fertilization, and pregnancy complicated with medical and surgical diseases. The study was approved by Hangzhou Women's Hospital [2022] Medical Ethics Review K (1)-09. Patient consent was not required because this was a retrospective study.

Comparison of Specimen Sources

Serum samples from 78 women with normal pregnancies were collected from May to August 2018, and serum samples from 4 pregnant women with T21 and 5 pregnant women with NTD were collected from September 2016 to June 2017. Eighty-seven serum specimens collected from women at 15 to 20 weeks and 6 days of pregnancy, which had previously undergone measurement of AFP, free β -hCG, and uE3 concentrations were randomly selected and briefly stored at -80°C . After thoroughly thawing and mixing, each specimen was divided into

six 250 µL fractions, placed into an ice box, and transported to the 6 laboratories at 2 to 8°C for testing.

Determination of Multiple of the Median Values for AFP, Free β-hCG, and uE3

We determined the multiple of the median (MoM) values for maternal weight and gestational age at testing to replace the concentrations of AFP, free β-hCG, and uE3 to reduce the deviation caused by gestational age and maternal weight. In brief, the MoM values were normalized to maternal weight and gestational age at testing [17] and subsequently adjusted according to the median equation of each laboratory.

The definition and calculation of MoM value was:

$$MoM = \frac{\text{Original Conj.}}{\text{Median}}$$

where Original Conj. was the original concentration values of AFP, free β-hCG, and uE3 and Median represented the median of the original concentration value of the corresponding indicator.

Quality Controls

All 6 instruments were maintained by trained users on a daily, weekly, and monthly basis, and the manufacturer provided unified quarterly maintenance to make each instrument reach the best state. Also, 3 different quality control serum specimens were conducted as part of internal quality control in the 6 laboratories. Each laboratory was required to participate in the EQA activities organized by the Clinical Laboratory Center of China Health Commission every year. In addition, all personnel were required to show proof of unified pre-job training and completed qualification certificates from health authorities.

Evaluation Standards

All laboratories used Life Cycle 4.0 software to calculate the risk values of T21 and NTD based on maternal age, gestational age at testing, and maternal weight. The cut-off values according to the 2010 China Ministry of Health Guidelines for maternal prenatal screening in the second trimester [18-19] were as follows: high risk (T21 ≥1: 270, T18 ≥1: 350, NTD AFP MoM ≥2.50) and low risk (T21 <1: 270, T18 <1: 350; NTD AFP MoM <2.50).

Prenatal Diagnosis

The indications were as follows: pregnant women over 35 years old with high risk of fetal chromosomal abnormalities screened before delivery and who had given birth to children with chromosome diseases; pregnant women with suspected

fetal chromosomal abnormalities by prenatal ultrasound examination; the husband and wife were carriers of chromosomal abnormalities; doctors considered it necessary to carry out prenatal diagnosis in other cases. Pregnant women at 17 to 20 weeks of pregnancy underwent ultrasound-guided amniocentesis with informed consent as follows. After placing of the amniotic fluid tank, a 20 G puncture needle was used to collect 30 mL of amniotic fluid from the abdomen and place it in a sterile special centrifuge tube for examination. It was then centrifuged at 2500 r/min for 10 min, cell microspheres were cultured (37°C, 5% CO₂), and cell growth was observed. Thirty mitotic phases were counted under the microscope, 5 karyotypes were analyzed, 400 G-banded chromosomes were produced, and karyotypes were analyzed according to the international nomenclature system of human genetics. The number of abnormal karyotypes was increased, and C and N bands were added if necessary [20]. Ultrasound diagnosis was performed in accordance with the guidelines for prenatal ultrasonography [21].

Statistical Analysis

R software (version 4.0.3; <https://www.R-project.org>) was used for drawing, statistical data analysis, and visualization. Lilliefors (1967) modified test (Shapiro-Wilk test) was used to determine whether the data was normally distributed. AFP and free β-hCG concentrations showed a skewed distribution; therefore, they were expressed as medians and percentiles (M [P_{2.5}, P_{97.5}]). Skewed data was compared with the laboratories that used the Kruskal-Wallis H (K) test or the Mann-Whitney U test. The linear correlation analysis of MoM values was done with Pearson linear correlation, whereas the high-risk positivity rate and the detection rate was analyzed using the paired-sample chi-squared test (McNemar test). The Kappa consistency test was used to analyze the accuracy and the reproducibility across the different laboratories. *P* values <0.05 indicated that the difference was statistically significant. The sensitivity, specificity, positive predictive value, negative predictive value, FPR, and false negative rate (FNR) were used to evaluate the screening efficacy of T21 and NTD. To reduce the cross-correlation of AFP, free β-hCG, and uE3 concentrations between T21 and NTD, the corresponding cases were excluded when the detection rate was calculated.

Results

Comparison of Different Laboratory Screening Results

The 6 laboratories analyzed 87 serum specimens. The lowest AFP concentration (30.52 [22.96-44.42] U/mL) was reported by laboratory F, and the difference was statistically significant (*P*<0.01) (Figure 1A). There were no statistically significant

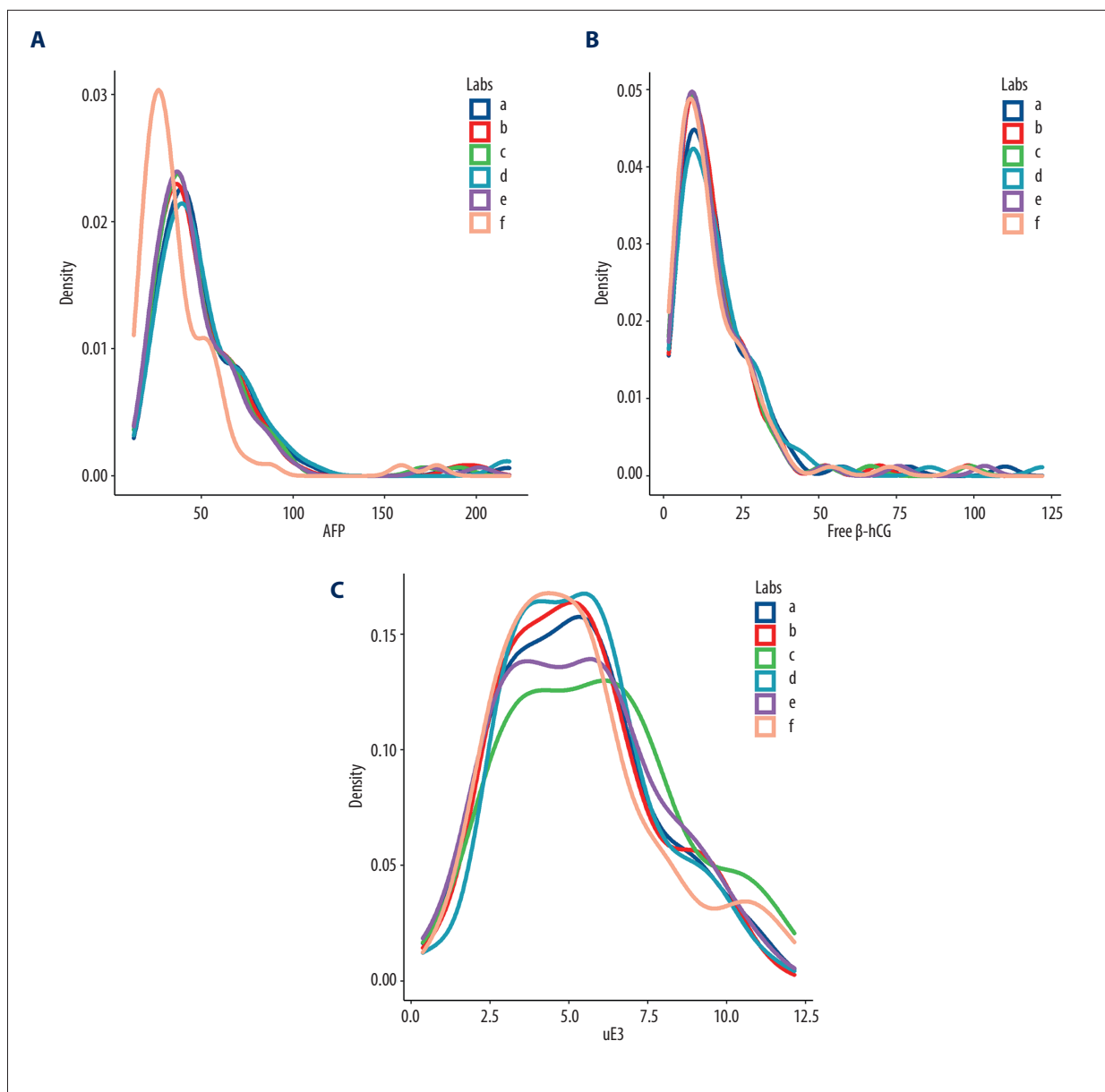


Figure 1. The levels of α -fetoprotein (AFP), free beta human chorionic gonadotropin (free β -hCG), and unconjugated estriol (uE3) in different screening laboratories. (A) AFP; (B) free β -hCG; (C) uE3. a – Laboratory A; b – Laboratory B; c – Laboratory C; d – Laboratory D; e – Laboratory E; f – Laboratory F.

differences in free β -hCG and uE3 concentrations among the 6 laboratories ($P > 0.05$) (Figure 1B, 1C). The MoM values for AFP, free β -hCG, and uE3 were also determined by the 6 laboratories (Figure 2). The MoM values for AFP and uE3, as detected by laboratory F, were the lowest, and the differences were statistically significant ($P < 0.001$; $P = 0.008$) (Figure 2A, 2C). However, there was no statistically significant difference in the free β -hCG MoM values among the 6 laboratories ($P > 0.05$) (Table 1, Figure 2B).

Correlation Analysis of the MoM Values for AFP, Free β -hCG, and uE3

The MoM values for AFP, free β -hCG, and uE3, as determined by the other 5 laboratories were comparable to those determined by laboratory A, with good linear correlation. The correlation coefficients for the 3 MoM values were all > 0.900 ($P < 0.05$) (Figures 3-5). Whereas the correlation coefficient for the uE3 MoM value was the lowest at 0.834, as detected by laboratory F (Table 2, Figure 5; Supplementary Tables 1-3).

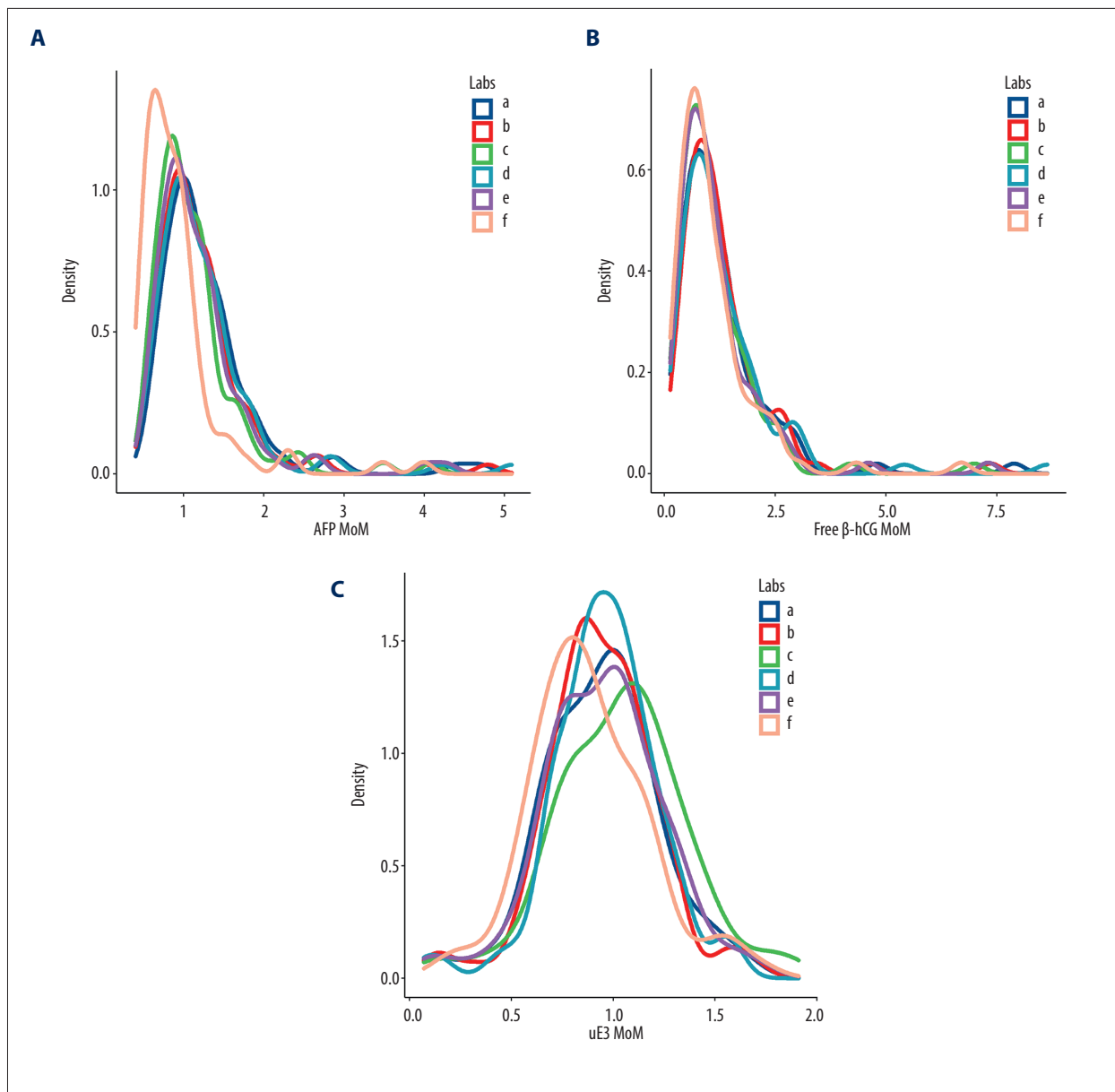


Figure 2. The level of α -fetoprotein (AFP), free beta human chorionic gonadotropin (free β -hCG), and unconjugated estriol (μ E3) multiple of the median (MoM) values in different screening laboratories. (A) AFP MoM; (B) free β -hCG MoM; (C) uE3 MoM. a – Laboratory A; b – Laboratory B; c – Laboratory C; d – Laboratory D; e – Laboratory E; f – Laboratory F.

Comparison of the T21 Positivity Rate and Detection Rate

Laboratories A and F detected the most cases of high-risk T21, with 6 (7.32%) and 7 (8.54%) cases, respectively, whereas the other 4 laboratories detected 4 cases (4.88%) each (Table 3). The McNemar paired chi-squared test showed that there was no significant difference between the T21 positivity rate and clinical diagnosis ($P>0.05$). The kappa consistency test was used to analyze the consistency of the different laboratory results. For laboratories B, C, D, and E, the experimental results had good consistency ($\text{kappa}=0.737$); however, for laboratories A

($\text{kappa}=0.575$) and F ($\text{kappa}=0.515$), the experimental results had mediocre consistency (Table 4).

Comparison of the NTD Positivity Rate and Detection Rate

McNemar's paired chi-squared test showed that laboratories C and F detected 1 case (1.20%) each of high-risk NTD, which was less than the 3 cases (3.61%) detected by the other 4 laboratories. However, there were no statistically significant differences in the positivity rate and detection rate of NTD ($P>0.05$) (Table 5). For laboratories A, B, D, and E, the experimental results

Table 1. Comparison of screening results in different screening laboratories.

Maternal markers	Lab A	Lab B	Lab C	Lab D	Lab E	Lab F	P value
AFP	43.19 (33.91, 63.04)	41.10 (32.00, 61.10)	41.30 (32.80, 60.20)	44.00 (34.30, 65.00)	40.90 (31.98, 58.64)	30.52 (22.96, 44.42)	<0.001
free β -hCG	14.17 (8.24, 22.35)	13.20 (8.11, 20.50)	12.40 (7.42, 20.80)	14.20 (7.92, 23.20)	12.63 (7.63, 21.04)	12.15 (6.97, 21.56)	0.673
uE3	5.22 (3.47, 6.65)	4.99 (3.49, 6.55)	5.75 (3.59, 7.47)	5.30 (3.53, 6.48)	5.07 (3.26, 7.02)	4.98 (3.29, 6.36)	0.633
AFP MoM	1.11 (0.92, 1.47)	1.05 (0.85, 1.38)	0.97 (0.81, 1.25)	1.08 (0.88, 1.41)	1.01 (0.82, 1.35)	0.78 (0.61, 1.02)	<0.001
Free β -hCG MoM	1.01 (0.67, 1.52)	1.06 (0.71, 1.56)	0.88 (0.59, 1.47)	0.99 (0.61, 1.58)	0.94 (0.61, 1.38)	0.84 (0.56, 1.33)	0.342
uE3 MoM	0.95 (0.75, 1.11)	0.94 (0.80, 1.08)	1.05 (0.81, 1.22)	0.95 (0.84, 1.11)	0.96 (0.75, 1.12)	0.84 (0.70, 1.08)	0.008

Lab – laboratory; AFP – α -fetoprotein; free β -hCG – free beta human chorionic gonadotropin; uE3 – unconjugated estriol; MoM – multiple of median.

had good consistency ($\kappa=0.738$); however, for laboratories C ($\kappa=0.320$) and F ($\kappa=0.320$), the experimental results had poor consistency (Table 6).

Comparison of the T21 and NTD Screening Efficacy

The highest FPR for T21 screening was detected by laboratory F (5.13%), followed by laboratory A (3.85%) and the other 4 laboratories (1.28%) (Table 7). For NTD, the lowest sensitivity of NTD screening were detected by laboratory C (20.00%); the highest FNR of NTD screening were detected by laboratory F (80.00%).

Discussion

EQA in Different Laboratory Detection Systems

The 6 laboratories in this study determined and calculated the risk of T21 and NTD using maternal serum specimens in which the concentrations of AFP, free β -hCG, and uE3 were measured. EQA can objectively and retrospectively compare the results of different laboratories through proficiency testing, which is maintained by external institutions. The main purpose of EQA is to establish a robust comparability between laboratories and methods and to reach agreement with established reference values [22].

The automatic time-resolved fluorescence analyzer used in this study employed a standardized protocol of loading specimens and reagents uniformly, incubating and shaking specimens for fixed periods of time, washing plates consistently,

and adding equivalent volumes of enhancement reagent [23]. Yee et al reported that it is best to define as many parameters as possible, such as the pre-analytic process, specimen type, platform type, and test components, as well as positive and negative controls [24]. A previous study showed the 5420 pregnant women were enrolled for triple test investigations in the second trimester, which can be used as reference population – specific median values in India [25]. EQA is often used in the determination of biomarker concentrations, which is critical in the diagnosis and prognosis of certain diseases, such as T21 and NTD, as well as in the prediction of treatment outcomes [26,27]. Our results revealed that the AFP concentration and AFP MoM value, as determined by laboratory F, were the lowest among the 6 laboratories, and the difference was statistically significant. This might be related to the newer version of the instrument system tested in the laboratory and the relatively small daily sample size of the laboratory.

Comparison of AFP, Free β -hCG, and uE3

Although there is an international reference value for AFP, the AFP detection method across different laboratories is different, which renders the reference value useless in analysis. As such, laboratories should use quality control serum specimens that are not included in commercially-available EQA kits [28]. The simultaneous quantification of AFP and free β -hCG levels is common in the screening of T21 in the second trimester of pregnancy, and its efficacy is assessed by the FPR. Any bias in MoM values may result in the incorrect application of median equations, and regular monitoring of these indicators may be essential for quality control in prenatal screening [29]. Mannings et al reported that the AFP concentration,

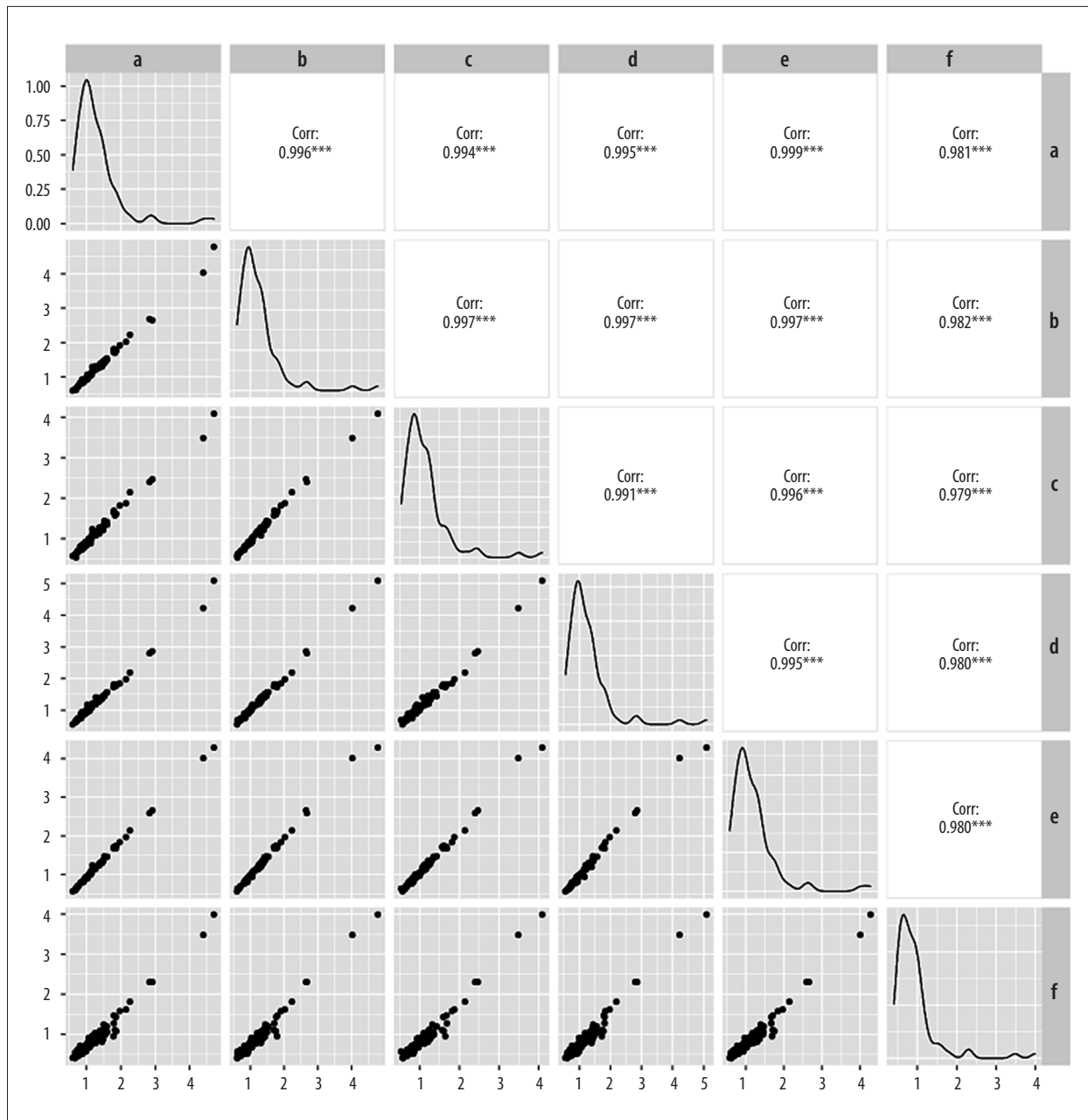


Figure 3. Linear correlation analysis of α -fetoprotein (AFP) multiple of the median (MoM) value. a – Laboratory A; b – Laboratory B; c – Laboratory C; d – Laboratory D; e – Laboratory E; f – Laboratory F.

as measured by the automatic time-resolved fluorescence analyzer, can indeed lower the AFP concentration due to complement interference, which can increase the calculated risk of T21 [30]. Our results showed that there was no statistically significant difference in the free β -hCG MoM value determined by the 6 laboratories. However, it is important to mention that laboratories implementing the new platform, which employs the new assay, will need to determine new reference values (medians), although this is not expected to have a marked effect on the maternal prenatal screening of T21 [31].

There was no statistically significant difference in the uE3 concentrations measured by the 6 laboratories. However, the uE3 MoM value determined by laboratory F was the lowest, and the difference was statistically significant, indicating that the MoM value, which was normalized against maternal weight and gestational age at testing, was more accurate than the concentration. Thus, it is necessary to simultaneously monitor 2 or more assays, which can be easily accomplished by monitoring the MoM values for each assay performed in the maternal prenatal screening of T21. The monitoring of MoM

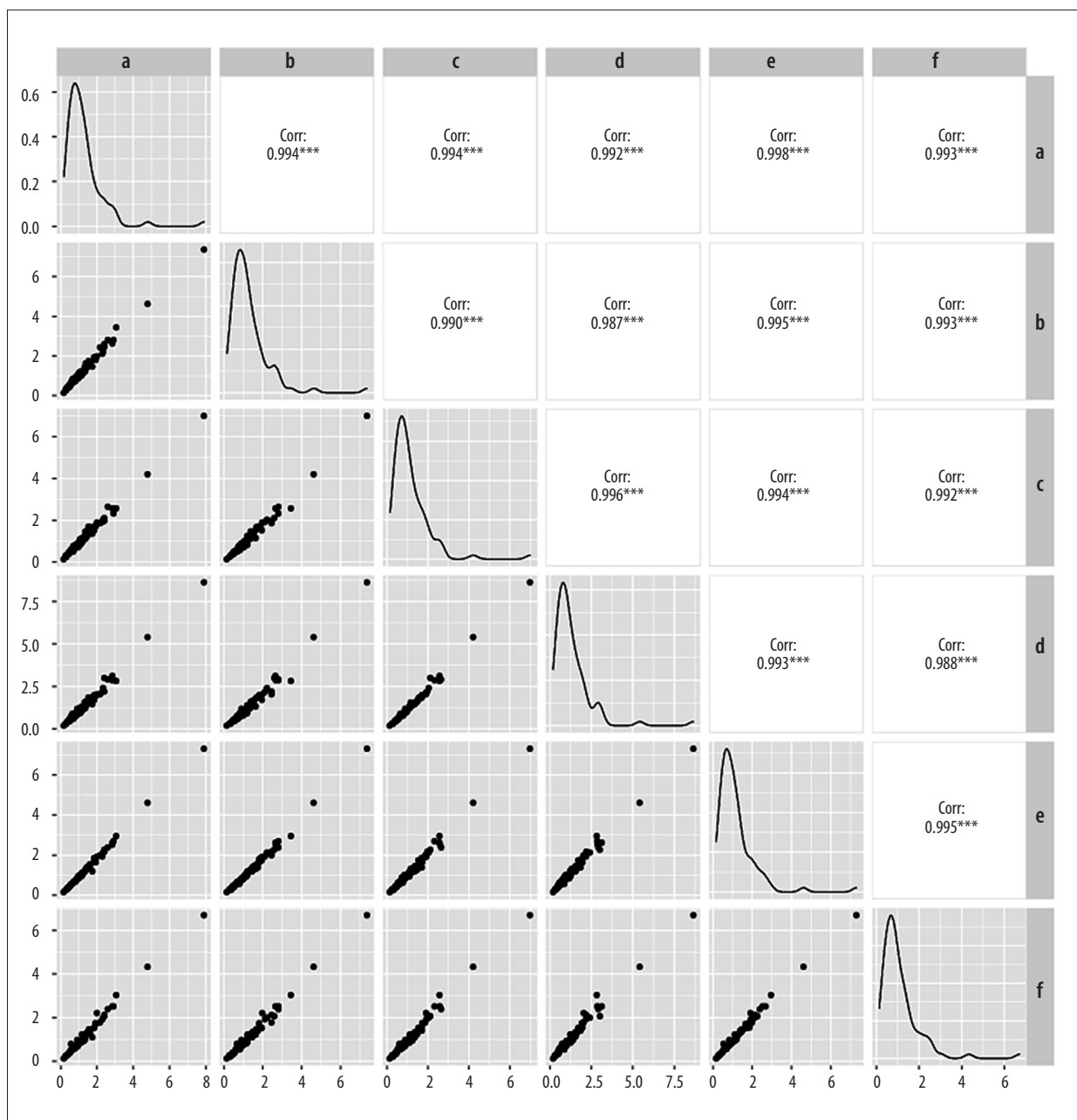


Figure 4. Linear correlation analysis of free beta human chorionic gonadotropin (free β -hCG) multiple of the median (MoM) value. a – Laboratory A; b – Laboratory B; c – Laboratory C; d – Laboratory D; e – Laboratory E; f – Laboratory F.

values can identify factors that inappropriately alter critical variables. For instance, the integration of MoM values can increase the detection rate when interpreting the test results of T21 screening [32,33].

For specific individuals, the use of localized parameter calculations can be more accurate because these calculations are significant in maternal prenatal screening [34]. Nix et al reported that a 10% bias in a single marker can lead to a 1% to 2% increase in the FPR and that quality control measures should

be established to control the magnitude of the bias associated with the MoM value or the magnitude of the bias associated with the calculation of patient-specific risks [35]. The uE3 MoM value was 1.05, as determined by laboratory C, which was higher than that of the other 5 laboratories. This might have been related to the fact that the uE3 assay was just performed and the median equation for uE3 was not yet adjusted by the 6 laboratories.

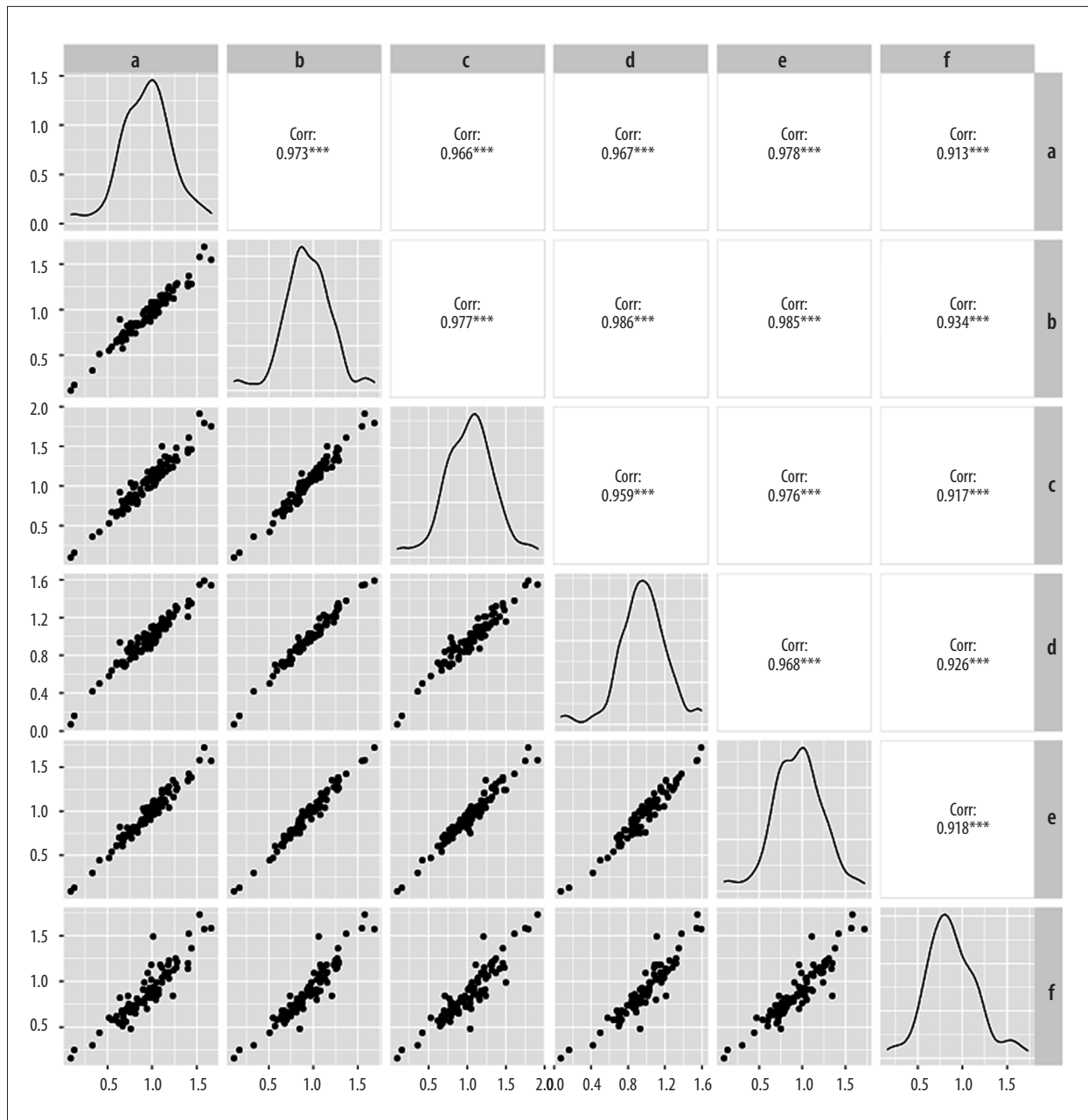


Figure 5. Linear correlation analysis of unconjugated estriol (uE3) multiple of median (MoM) value. a – Laboratory A; b – Laboratory B; c – Laboratory C; d – Laboratory D; e – Laboratory E; f – Laboratory F.

Comparison of T21 and NTD Diagnosed Cases

The MoM values for AFP, free β -hCG, and uE3 determined by the 5 laboratories were comparable with those determined by laboratory A, with good linear correlation. The correlation coefficient was >0.900 , and the difference was statistically significant, indicating that the test results met the clinical requirements. Laboratories A and F detected the most cases of high-risk T21. The positivity rate of T21, as determined by laboratories A and F, was higher than that of the other 4

laboratories, whereas the positivity rate of NTD, as determined by laboratories C and F, was lower than that of the other laboratories, although the differences were not statistically significant. Furthermore, the results of the kappa consistency test revealed that the findings of most laboratories had good consistency. However, the accuracy and the reproducibility of the maternal prenatal screening for T21 in the first trimester in China should be further improved [36].

Table 2. Correlation analysis of screening results between 5 screening laboratories and Laboratory A.

Laboratory/screening indicators	R value	The regression equation	P value	R ²
Lab B				
AFP MoM	0.996	Y=-0.016+0.966X	<0.001	0.993
free β-hCG MoM	0.994	Y=0.080+0.951X	<0.001	0.988
uE3 MoM	0.973	Y=0.078+0.915X	<0.001	0.946
Lab C				
AFP MoM	0.994	Y=-0.016+0.966X	<0.001	0.987
free β-hCG MoM	0.994	Y=0.080+0.951X	<0.001	0.988
uE3 MoM	0.966	Y=0.078+0.915X	<0.001	0.932
Lab D				
AFP MoM	0.995	Y=-0.064+1.020X	<0.001	0.992
free β-hCG MoM	0.992	Y=-0.065+1.080X	<0.001	0.994
uE3 MoM	0.967	Y=0.138+0.877X	<0.001	0.935
Lab E				
AFP MoM	0.999	Y=0.010+0.916X	<0.001	0.998
free β-hCG MoM	0.998	Y=-0.016+0.929X	<0.001	0.996
uE3 MoM	0.978	Y=0.008+1.000X	<0.001	0.957
Lab F				
AFP MoM	0.981	Y=-0.136+0.824X	<0.001	0.962
free β-hCG MoM	0.993	Y=-0.015+0.879X	<0.001	0.986
uE3 MoM	0.913	Y=0.019+0.920X	<0.001	0.834

Lab – laboratory; AFP – α-fetoprotein; free β-hCG – free beta human chorionic gonadotropin; uE3 – unconjugated estriol; MoM – multiple of median.

Table 3. Comparison of positive rate and detection rate of trisomy 21 in 6 different laboratories.

Laboratory	N	T21 risk value distribution			T21 detection situation		
		Low risk	High risk	PR (%)	Low risk	High risk	DR (%)
Lab A	82	76	6	7.32	1	3	75.00
Lab B	82	78	4	4.88	1	3	75.00
Lab C	82	78	4	4.88	1	3	75.00
Lab D	82	78	4	4.88	1	3	75.00
Lab E	82	78	4	4.88	1	3	75.00
Lab F	82	75	7	8.54	1	3	75.00

Lab – laboratory; T21 – trisomy 21; PR – positive rate; DR – detection rate.

It is not advisable to improve the accuracy and the reproducibility of test results by requiring that all laboratories participate in EQA programs, although such programs may quickly identify issues with the procedures or the methods of laboratories. Inaccurate and non-reproducible test results should be promptly addressed by appropriate measures [37,38]. For instance, EQA monitoring can identify issues that are not readily

detected with conventional quality control measures. There are different calibration system errors and individual differences of operators in different laboratories to facilitate the homogenization management of prenatal screening reports among different laboratories. As the quality control management center in the region, we compared the results of 6 laboratories using the same samples and laboratory equipment.

Table 4. Comparison of T21 screening effect in 6 different laboratories.

Laboratory	High/Low risk	T21 diagnostic result		Total	P value	Kappa value
		Normal	T21			
Lab A	Low risk	75	1	76	0.625	0.575
	High risk	3	3	6		
Lab B	Low risk	77	1	78	1.000	0.737
	High risk	1	3	4		
Lab C	Low risk	77	1	78	1.000	0.737
	High risk	1	3	4		
Lab D	Low risk	77	1	78	1.000	0.737
	High risk	1	3	4		
Lab E	Low risk	77	1	78	1.000	0.737
	High risk	1	3	4		
Lab F	Low risk	74	1	75	0.375	0.515
	High risk	4	3	7		

Lab – laboratory; T21 – trisomy 21.

Table 5. Comparison of positive rate and detection rate of neural tube defects in 6 different laboratories.

Laboratory	N	NTD risk value distribution			NTD detection situation				
		Low risk	High risk	PR (%)	P value	Low risk	High risk	DR (%)	P value
Lab A	83	80	3	3.61	0.799	2	3	75.00	0.529
Lab B	83	80	3	3.61		2	3	75.00	
Lab C	83	82	1	1.20		4	1	25.00	
Lab D	83	80	3	3.61		2	3	75.00	
Lab E	83	80	3	3.61		2	3	75.00	
Lab F	83	82	1	1.20		4	1	25.00	

Lab – laboratory; NTD – neural tube defects; PR – positive rate; DR – detection rate.

Limitations

This study had some limitations. It was limited to 87 serum samples, including only 4 confirmed cases of T21 and 5 confirmed cases of NTD, while cases of T18 were not included. There may be bias in the comparison of AFP, free β -hCG, and uE3 among the 6 instruments based on such a small sample. Further validation will be done by increasing the sample size in later studies.

Conclusions

In summary, the MoM values of AFP, free β -hCG, and uE3, as determined by the 5 laboratories, were comparable with those determined by laboratory A, with good linear correlation. The

correlation coefficient was >0.900 , and the difference was statistically significant, indicating that the test results met the clinical requirements, except that the kappa consistency of a few laboratories was poor. After improving the quality control of various factors that can affect test results, the median equation calibration work should be performed according to the comparison results to reduce all errors to their lowest levels. The values of AFP, free β -hCG, and uE3 measured in 5 laboratories were comparable to those measured in laboratory A and had good linear correlation. The results of prenatal screening for pregnant women with T21 and NTD could meet the clinical requirements of their institutions.

Table 6. Comparison of neural tube defect screening effect in 6 different laboratories.

Lab	Low/High risk	NTD diagnostic result		Total	P value	Kappa value
		Normal	NTD			
Lab A	Low risk	78	2	80	0.500	0.738
	High risk	0	3	3		
Lab B	Low risk	78	2	80	0.500	0.738
	High risk	0	3	3		
Lab C	Low risk	78	4	82	0.125	0.320
	High risk	0	1	1		
Lab D	Low risk	78	2	80	0.500	0.738
	High risk	0	3	3		
Lab E	Low risk	78	2	80	0.500	0.738
	High risk	0	3	3		
Lab F	Low risk	78	4	82	0.125	0.320
	High risk	0	1	1		

Lab – laboratory; NTD – neural tube defects; PR – positive rate; DR – detection rate.

Table 7. Comparison of screening efficiency of trisomy 21 and neural tube defects in different laboratories (%).

Lab	Sensitivity	Specificity	PPV	NPV	FPR	FNR
T21						
Lab A	75.00	96.15	50.00	98.68	3.85	25.00
Lab B	75.00	98.72	75.00	98.72	1.28	25.00
Lab C	75.00	98.72	75.00	98.72	1.28	25.00
Lab D	75.00	98.72	75.00	98.72	1.28	25.00
Lab E	75.00	98.72	75.00	98.72	1.28	25.00
Lab F	75.00	94.87	42.86	98.67	5.13	25.00
NTD						
Lab A	60.00	100.00	100.00	97.50	0.00	40.00
Lab B	60.00	100.00	100.00	97.50	0.00	40.00
Lab C	20.00	100.00	100.00	95.12	0.00	80.00
Lab D	60.00	100.00	100.00	97.50	0.00	40.00
Lab E	60.00	100.00	100.00	97.50	0.00	40.00
Lab F	20.00	100.00	100.00	95.12	0.00	80.00

Lab – laboratory; PPV – positive predictive value; NPV – negative predictive value; FPR – false positive rate; FNR – false negative rate; NTD – neural tube defects; T21 – trisomy 21.

Acknowledgments

The authors are grateful to all participants and contributors. We also thank International Science Editing (<http://www.internationalscienceediting.com>) for editing this manuscript.

Declaration of Figures’ Authenticity

All figures submitted have been created by the authors, who confirm that the images are original with no duplication and have not been previously published in whole or in part

Supplementary Table 1. Correlation analysis of α -fetoprotein (AFP) multiple of the median (MoM) in 6 laboratories.

Laboratory		Lab A	Lab B	Lab C	Lab D	Lab E	Lab F
Lab A	<i>Pearson correlation</i>	1	0.996**	0.994**	0.995**	0.999**	0.981**
	<i>P</i>		<0.001	<0.001	<0.001	<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab B	<i>Pearson correlation</i>	0.996**	1	0.997**	0.997**	0.997**	0.982**
	<i>P</i>	<0.001		<0.001	<0.001	<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab C	<i>Pearson correlation</i>	0.994**	0.997**	1	0.991**	0.996**	0.979**
	<i>P</i>	<0.001	<0.001		<0.001	<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab D	<i>Pearson correlation</i>	0.995**	0.997**	0.991**	1	0.995**	0.980**
	<i>P</i>	<0.001	<0.001	<0.001		<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab E	<i>Pearson correlation</i>	0.999**	0.997**	0.996**	0.995**	1	0.980**
	<i>P</i>	<0.001	<0.001	<0.001	<0.001		<0.001
	<i>N</i>	87	87	87	87	87	87
Lab F	<i>Pearson correlation</i>	0.981**	0.982**	0.979**	0.980**	0.980**	1
	<i>P</i>	<0.001	<0.001	<0.001	<0.001	<0.001	
	<i>N</i>	87	87	87	87	87	87

** Correlation is significant at a 0.01 level (2-tailed).

Supplementary Table 2. Correlation analysis of free beta human chorionic gonadotropin (free β -hCG) multiple of the median (MoM) value in 6 laboratories.

Laboratory		Lab A	Lab B	Lab C	Lab D	Lab E	Lab F
Lab A	<i>Pearson correlation</i>	1	0.994**	0.994**	0.992**	0.998**	0.993**
	<i>P</i>		<0.001	<0.001	<0.001	<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab B	<i>Pearson correlation</i>	0.994**	1	0.990**	0.987**	0.995**	0.993**
	<i>P</i>	<0.001		<0.001	<0.001	<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab C	<i>Pearson correlation</i>	0.994**	0.990**	1	0.996**	0.994**	0.992**
	<i>P</i>	<0.001	<0.001		<0.001	<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab D	<i>Pearson correlation</i>	0.992**	0.987**	0.996**	1	0.993**	0.988**
	<i>P</i>	<0.001	<0.001	<0.001		<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab E	<i>Pearson correlation</i>	0.998**	0.995**	0.994**	0.993**	1	0.995**
	<i>P</i>	<0.001	<0.001	<0.001	<0.001		<0.001
	<i>N</i>	87	87	87	87	87	87
Lab F	<i>Pearson correlation</i>	0.993**	0.993**	0.992**	0.988**	0.995**	1
	<i>P</i>	<0.001	<0.001	<0.001	<0.001	<0.001	
	<i>N</i>	87	87	87	87	87	87

** Correlation is significant at a 0.01 level (2-tailed).

Supplementary Table 3. Correlation analysis of unconjugated estriol (uE3) multiple of median (MoM) value in 6 laboratories.

Laboratory		Lab A	Lab B	Lab C	Lab D	Lab E	Lab F
Lab A	<i>Pearson correlation</i>	1	0.973**	0.966**	0.967**	0.978**	0.913**
	<i>P</i>		<0.001	<0.001	<0.001	<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab B	<i>Pearson correlation</i>	0.973**	1	0.977**	0.986**	0.985**	0.934**
	<i>P</i>	<0.001		<0.001	<0.001	<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab C	<i>Pearson correlation</i>	0.966**	0.977**	1	0.959**	0.976**	0.917**
	<i>P</i>	<0.001	<0.001		<0.001	<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab D	<i>Pearson correlation</i>	0.967**	0.986**	0.959**	1	0.968**	0.926**
	<i>P</i>	<0.001	<0.001	<0.001		<0.001	<0.001
	<i>N</i>	87	87	87	87	87	87
Lab E	<i>Pearson correlation</i>	0.978**	0.985**	0.976**	0.968**	1	0.918**
	<i>P</i>	<0.001	<0.001	<0.001	<0.001		<0.001
	<i>N</i>	87	87	87	87	87	87
Lab F	<i>Pearson correlation</i>	0.913**	0.934**	0.917**	0.926**	0.918**	1
	<i>P</i>	<0.001	<0.001	<0.001	<0.001	<0.001	
	<i>N</i>	87	87	87	87	87	87

** Correlation is significant at a 0.01 level (2-tailed).

References:

- Chitayat D, Langlois S, Wilson RD. No. 261-prenatal screening for fetal aneuploidy in singleton pregnancies. *J Obstet Gynaecol Can.* 2017;39(9):e380-94
- Tu S, Rosenthal M, Wang D, et al. Performance of prenatal screening using maternal serum and ultrasound markers for Down syndrome in Chinese women: A systematic review and meta-analysis. *Br J Obstet Gynaecol.* 2016;123(Suppl. 3):12-22
- Kaur G, Srivastav J, Kaur A, et al. Maternal serum second trimester screening for chromosomal disorders and neural tube defects in a government hospital of North India. *Prenat Diagn.* 2012;32(12):1192-96
- Allred SK, Deeks JJ, Guo B, et al. Second trimester serum tests for Down's Syndrome screening. *Cochrane Database Syst Rev.* 2012(6):D9925
- Chard T. Biochemistry and endocrinology of the Down's syndrome pregnancy. *Ann NY Acad Sci* 1991;626:580-96
- Wald NJ, Cuckle H, Brock JH, et al. Maternal serum-alpha-fetoprotein measurement in antenatal screening for anencephaly and spina bifida in early pregnancy. Report of U.K. collaborative study on alpha-fetoprotein in relation to neural-tube defects. *Lancet.* 1977;1(8026):1323-32
- Tsou AY, Bulova P, Capone G, et al. Medical care of adults with Down syndrome: A clinical guideline. *JAMA.* 2020;324(15):1543-56
- Douglas WR, Van Mieghem T, Langlois S, Church P. Guideline No. 410: Prevention, screening, diagnosis, and pregnancy management for fetal neural tube defects. *J Obstet Gynaecol Can.* 2021;43(1):124-39
- Lou S, Mikkelsen L, Hvidman L, et al. Does screening for Down's syndrome cause anxiety in pregnant women? A systematic review. *Acta Obstet Gynecol Scand.* 2015;94(1):15-27
- Johnston J, Farrell RM, Parens E. Supporting women's autonomy in prenatal testing. *N Engl J Med.* 2017;377(6):505-7
- Dane AC, Peterson M, Miller YD. Talking points: Women's information needs for informed decision-making about noninvasive prenatal testing for Down syndrome. *J Genet Couns.* 2018;27:1258-64
- Than NG, Papp Z. Ethical issues in genetic counseling. *Best Pract Res Clin Obstet Gynaecol.* 2017;43:32-49
- Stavelin A, Albe X, Meijer P, et al. An overview of the European organization for external quality assurance providers in laboratory medicine (EQALM). *Biochem Med (Zagreb).* 2017;27:30-36
- Ding T, Bergeron M, Seely P, et al. Compatibility of stabilized whole blood products with CD4 technologies and their suitability for quality assessment programs. *PLoS One.* 2014;9(8):e103391
- Melese M, Jerene D, Alem G, et al. Decentralization of Acid Fast Bacilli (AFB) external quality assurance using blind rechecking for sputum smear microscopy in Ethiopia. *PLoS One.* 2011;6(11):e2151366
- Verderio P, Pizzamiglio S, Ciniselli CM. Methodological and statistical issues in developing an External Quality Assessment scheme in laboratory medicine: Focus on biomarker research. *N Biotechnol.* 2019;52:54-59
- Chen Y, Xie Z, Wang X, et al. A risk model of prenatal screening markers in first trimester for predicting hypertensive disorders of pregnancy. *EPMA J* 2020;11(3): 343-53
- Jiang T, Ding J, Zhang XQ, et al. Analysis of Down syndrome failed to be diagnosed after prenatal screening: A multicenter study. *Medicine (Baltimore).* 2017;96(24):e7166
- Miao ZY, Liu X, Shi TK, et al. First trimester, second trimester, and integrated screening for Down's syndrome in China. *J Med Screen.* 2012;19(2):68-71
- Ordulu Z, Wong K-E, Currell B-B, et al. Describing sequencing results of structural chromosome rearrangements with a suggested next-generation cytogenetic nomenclature. *Am J Hum Genet.* 2014;94(5):695-709
- Reddy UM, Abuhamad AZ, Levine D, Saade GR. Fetal imaging: Executive summary of a Joint Eunice Kennedy Shriver National Institute of Child Health and Human Development, Society for Maternal-Fetal Medicine, American Institute of Ultrasound in Medicine, American College of Obstetricians and Gynecologists, American College of Radiology, Society for Pediatric Radiology, and Society of Radiologists in Ultrasound Fetal Imaging Workshop. *Am J Obstet Gynecol.* 2014;210(5):387-97

22. England JM, Rowan RM, van Assendelft OW, et al. Guidelines for organisation and management of external quality assessment using proficiency testing. Expert Panel on Cytometry of the International Council for Standardization in Haematology. *Int J Hematol.* 1998;68(1):45-52
23. Hemmila I, Dakubu S, Mikkala VM, et al. Europium as a label in time-resolved immunofluorometric assays. *Anal Biochem.* 1984;137(2):335-43
24. Yee LM, Lively TG, McShane LM. Biomarkers in early-phase trials: Fundamental issues. *Bioanalysis.* 2018;10:933-44
25. Kaur G, Srivastav J, Sharma S, et al. Maternal serum median levels of alpha-fetoprotein, human chorionic gonadotropin & unconjugated estriol in second trimester in pregnant women from north-west India. *Indian J Med Res.* 2013;138(1):83-88
26. Tsai JR, Liu PL, Chen YH, et al. Ginkgo biloba extract decreases non-small cell lung cancer cell migration by downregulating metastasis-associated factor heat-shock protein 27. *Plos One.* 2014;9(3):e91331
27. de Leeuw RA, van der Horst S, de Soet AM, et al. Digital vs face-to-face information provision in patient counselling for prenatal screening: A non-inferiority randomized controlled trial. *Prenat Diagn.* 2019;39(6):456-63
28. Oremek GM, Oertl A, Bertsch T, et al. Alpha-1-Fetoprotein (AFP): International proficiency study with different test systems. *Clin Lab.* 2011;57:669-75
29. Li Y, Zhang X, Hong D, et al. Significance of data analysis in the quality control of prenatal screening for Down syndrome. *Biomed Rep.* 2018;8(5):447-53
30. Mannings L, Trow S, Newman J, et al. Interference in the auto DELFIA(R) hAFP immunoassay and effect on second-trimester Down's syndrome screening. *Ann Clin Biochem.* 2011;48(Pt 5):438-40
31. Palomaki GE, Lambert-Messerlian G. Down syndrome screening: Suitability of a WHO 5 standardized total hCG assay. *Clin Biochem.* 2014;47(7-8):629-31
32. Zhou Y, Du Y, Zhang B, Wang L. Integrating multiple of the median values of serological markers with the risk cut-off value in Down syndrome screening. *Biosci Trends.* 2018;12(6):613-19
33. Knight GJ. Quality assessment of a prenatal screening program. *Early Hum Dev.* 1996;47(Suppl.):S49-53
34. Hu ZM, Luo LL, Li L, et al. Indigenization of the median of markers for Down syndrome screening based on statistical analysis of medical big data. *Taiwan J Obstet Gynecol.* 2020;59(4):556-64
35. Nix B, Wright D, Baker A. The impact of bias in MoM values on patient risk and screening performance for Down syndrome. *Prenat Diagn.* 2007;27:840-45
36. He F, Wang W, Zhong K, et al. Analysis of external quality assessment of maternal serum prenatal screening for Down syndrome in the first trimester in China. *Clin Lab.* 2021;67(6):201004
37. Plebani M. Errors in laboratory medicine and patient safety: the road ahead. *Clin Chem Lab Med.* 2007;45:700-7
38. James D, Ames D, Lopez B, et al. External quality assessment: Best practice. *J Clin Pathol.* 2014;67:651-55