

Research Article

Didem Derici Yıldırım*, Şenay Balcı Fidancı, Merve Türkegün Şengül, M. Burak Y. Çimen and Lülüfer Tamer



Establishing continuous reference intervals for liver function analytes using fractional polynomial regression

<https://doi.org/10.1515/tjb-2024-0086>

Received April 16, 2024; accepted January 28, 2025;

published online February 20, 2025

Abstract

Objectives: Reference intervals (RIs) play a crucial role in contemporary healthcare by providing critical support for clinical decision-making. RI calculation methods have certain limitations. They cannot capture age-related variations in laboratory values and non-linear relations between variables. Therefore, alternative methods that preserve the continuous type of age and consider non-linear relations must be explored. This study aimed to suggest an alternative method, namely, fractional polynomial (FP) regression models, to overcome these limitations and estimate the reference limits based on this method for ALT and AST markers.

Methods: This retrospective study is conducted on a large hospital database to establish continuous age and sex-specific RIs for ALT and AST in a Turkish population. The study included 4,386 and 2,834 individuals for ALT and AST, respectively, aged 18–92 years. FP models and a non-parametric approach are used to calculate RIs.

Results: For AST in females, a linear model indicated the relationship between age and AST in females ($p=0.673$).

However, more complex models ($m_0=4$) were explored for female ALT values. After several comparisons, the FP3 model was chosen as the best fit ($p=0.098$). For male ALT and AST, FP analysis showed no significant age-related effects ($p=0.679$ and $p=0.507$, respectively). Therefore, traditional non-parametric methods were used to calculate RIs.

Conclusions: Compared with discrete RIs, continuous RIs are more accurate. Laboratories should determine RIs that reflect their population. Future research could explore the application of FP models in the pediatric population.

Keywords: continuous reference intervals; fractional polynomial regression; ALT; AST; adults

Introduction

Reference intervals (RIs) are crucial in contemporary healthcare as they provide critical support for clinical decision-making [1]. The establishment of RIs is conducted in accordance with the protocols set out by the Clinical Laboratory Standards Institute (CLSI) and the International Federation of Clinical Chemistry (IFCC). They are defined as the middle 95 % range of laboratory results obtained from a population of healthy individuals [2]. RIs are used to interpret a patient's laboratory test results [3–5], particularly for liver function tests, such as alanine aminotransferase (ALT) and aspartate aminotransferase (AST). These tests are used to diagnose liver diseases and monitor treatment efficacy [6]. Inaccurate RIs can lead to misdiagnosis, ineffective treatment, and compromised patient outcomes. Therefore, improving RI determination methods is significant in contemporary healthcare [7, 8]. Particularly, the non-linear relationship between factors such as age and ALT/AST levels should be considered.

Traditional RI determination methods involve dividing the population into subgroups based on age, sex, and other factors and have several limitations [2, 9–11]. Particularly, these methods may not consider the non-linear relationship between these factors and ALT and AST levels. For example,

*Corresponding author: Didem Derici Yıldırım, PhD, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Mersin University, Mersin, Türkiye, E-mail: didemderici@mersin.edu.tr. <https://orcid.org/0000-0001-7709-6133>

Şenay Balcı Fidancı, M. Burak Y. Çimen and Lülüfer Tamer, Faculty of Medicine, Department of Medical Biochemistry, Mersin University, Mersin, Türkiye, E-mail: sbfidanci@hotmail.com (Ş. Balcı Fidancı), mybcimen@hotmail.com (M.B.Y. Çimen), lutamer@yahoo.com (L. Tamer). <https://orcid.org/0000-0002-7498-604X> (Ş. Balcı Fidancı). <https://orcid.org/0000-0002-1274-3499> (M.B.Y. Çimen). <https://orcid.org/0000-0002-0997-0260> (L. Tamer)

Merve Türkegün Şengül, Faculty of Medicine, Department of Biostatistics, Alanya Alaaddin Keykubat University, Antalya, Türkiye, E-mail: merveturkegun@gmail.com. <https://orcid.org/0000-0002-4405-521X>

categorizing age into discrete age groups may not accurately capture age-related variations in laboratory values. Consequently, lower and upper age limits may exhibit significant disparities when transitioning between two age groups, resulting in inaccurate medical diagnoses. Additionally, categorizing continuous predictors including age causes information loss, limiting RI accuracy [12–18]. To overcome these limitations and improve the accuracy of RIs for ALT and AST markers, alternative methods should be explored, such as fractional polynomial (FP) regression models [13, 17, 18].

FP regression models are statistical models that address the limitations of traditional RI determination methods [19, 20]. They are useful when variables are non-linear. They preserve the structure of continuous variables, such as age, by capturing non-linear relationships accurately. FP models are flexible and can capture non-linear relationships and thus are suitable for calculating RIs with continuous age variables including ALT and AST markers [21, 22].

This study aimed to (1) introduce and suggest FP models for RI calculation with continuous age and (2) estimate the reference limits based on this method for ALT and AST markers.

Materials and methods

Data collection

This retrospective study is conducted on a large hospital database to establish continuous age and sex-specific RIs for liver function values (ALT, AST) in a Turkish population. This study included 5,306 and 3,184 individuals for ALT and AST, respectively, who had undergone at least one liver function test between January 1, 2017, and June 30, 2023. Data of individuals aged 18–92 who applied to a single outpatient clinic for routine examination were obtained. Duplicate results from the same test on the same patient number were eliminated. We eliminated 235 duplicate records for ALT and 175 duplicate records for AST. These individuals applied to the outpatient clinic once and had a laboratory examination. Next, we excluded 1,020 outlier values for ALT and 342 outlier values for AST. Thus, the data consisted of the hospital population, with individuals who were most probably healthy. Finally, we included 4,051 and 2,667 individuals in the study for ALT and AST, respectively.

Age was measured in years. AST (U/L) and ALT (U/L) levels were examined using the enzymatic colorimetric method (AU5800, Beckman Coulter, CA, USA). The laboratory reference ranges for which the data from this study were obtained include AST values <38 U/L in males and <32 U/L in

females and ALT values <55 U/L. The gender variable was coded as *Male* or *Female*. In our laboratory, control materials are used for AST and ALT parameters at two different levels with 12-h intervals. These internal quality control results are evaluated within the framework of Westgard rules. For the external quality assessment, the RIQAS external quality control program is used for the relevant tests. The local research Ethics Committee approved the study on 06/03/2024 (Approval number: 2024/251). Informed consent was not required because the dataset consisted of de-identified secondary data released for research purposes.

Fractional polynomial regression

Royston et al. proposed FP functions, subsequently employed by Royston and Sauerbrei as a robust statistical approach for constructing medical research models [20, 23]. This method effectively addresses the issue of non-linear associations between the response variable and predictors, which is common in linear regression models. Categorizing continuous predictors may result in information loss [24] as it involves converting a continuous variable into discrete categories. Determining how many and which cut-points to employ in this process can be challenging as the underlying data distribution and desired level of granularity must be considered carefully [21, 25].

We expect one continuous predictor (X) and one response, and a suitable regression model is $\beta_1 X$. β_0 was subtracted to calculate easily. This approach can sufficiently describe the relationships. However, other models must be examined for potential enhancements. The simplest transformation model of the linear model is βX_1^p . Royston and Altman (1994) Developed FP models, which include transformations that are power functions X^p , X^{p_1} or $X^{p_1} + X^{p_2}$ for different p values of powers. The p values select from predefined limited set called $S = \{-2, -1, 0.5, 0, 0.5, 1, 2, 3\}$ and, $X^0 = \ln(X)$. FP functions with degrees of one (FP1) as defined as $FP1 = \beta X_1^p$. If $p=0$, $FP1 = \beta_1 \ln(X)$. FP1 contains eight models for eight different power values; the model that fits the data the best has the highest likelihood. For non-linearity test, $p=1$ vs. $p \neq 1$ is tested with a chi-square test with 1 degree of freedom. This test is based on the log-likelihood ratio ($-2\log L$). If the test statistic is not significant, the linear model is used; otherwise, an FP with a non-linear function is selected [21].

The more complicated function FP2 comes next after function FP1 as $FP2 = \beta X_1^{p_1} + \beta X_2^{p_2}$ ($p_1=1, p_2=2$). If $p_1=p_2=p$, it is called repeated powers, and $FP2 = \beta X_1^p + \beta X_2^p \ln(X)$. FP2 contains 36 power combinations, and the one with the highest likelihood is the best-fit model [21].

The best model is selected with closed testing procedures (called RA2) or sequential procedures [22, 26]. According to the results of the simulation study, the best model selection procedure is RA2 [26]. Although RA2 gives an approximate type 1 error value for each variable, the actual type 1 error rate could be higher than the nominal α value when the real relationship is linear in the sequential procedure [25]. RA2 procedure has three steps. First, the effect of X on the outcome is examined. Accordingly, the best FP2 is tested against the null model using the chi-square test with four degrees of freedom. If the test is not significant, the process is stopped. Thus, X is not effective. Otherwise, the process continues. Second, whether the relationship between X and the outcome is non-linear is determined. The best FP2 is tested against the linear model using the chi-square test with three degrees of freedom. If the test is not significant, then X and the outcome have a linear relationship, and the fit model is the linear model. Otherwise, the process continues. Third, the best FP model is determined. The best FP2 is tested against FP1 using the chi-square test with two degrees of freedom. If the test is not significant, then the best model is FP1 or FP2 otherwise [26].

Statistical analysis

Liver function analytes are presented as mean values with standard deviation (SD). Categorical variables are presented as numbers with percentages. A reference interval is created by assuming that calculated from a homogeneous population. Thus, elimination of outliers can be narrower reference intervals and the sample can be more homogeneous.

The Tukey method is a more sophisticated and common method that defined outliers as those points that fall below $Q1 - 1.5 \cdot IQR$ or above $Q3 + 1.5 \cdot IQR$, providing a clear, quantitative criterion for identifying data points that are significantly different from the bulk of the data in a set. The determination of sex partitions was conducted using Harris and Boyd's statistical test. This test was employed to assess whether each group exhibits statistically significant differences that justify their classification into distinct categories. Age-related RIs are calculated following the steps below:

Step 1: Fit the mean function

If the distribution of the measurements (the variable for which to establish a reference interval) shows skewness at different levels of age, the measurements are transformed logarithmically or using a Box-Cox power transformation. Post-transformation observations are represented by an FP model with age as the independent variable as proposed by

Altman and Chitty [27]. Royston and Wright recommend the use of FPs for modeling the mean or SD curve if a quartic or quantic polynomial is required for an adequate fit to the data [28].

A first-degree FP is raised to the power p , which is defined as one of the following: $\{x^{-2}, x^{-1}, x^{0.5}, x^0, x^{0.5}, x^1, x^2, x^3\}$. The researcher determines which equation fits best with the data.

Step 2: Study the residuals from the mean fit

The regression model presented in this study estimates the average value of the converted data based on the age variable denoted as *mean (age)*. The residuals of the regression model are computed. Altman's model utilizes weighted polynomial regression to estimate absolute residuals, which are then multiplied by $\sqrt{\pi/2}$ as a function of age [27].

Step 3: Fit an SD function

The RI for each age within the observed range is determined by calculating the mean age and adding or subtracting z times the SD of age. The RI with a 95 % confidence level is determined by the value of z , that is, 1.96. If the data were first subjected to a transformation in step 1, then the subsequent step involves back-transforming the findings to the original scale.

Step 4: Calculate Z-scores

After deriving the regression equations for the mean and SD from the study group, the z -scores are calculated. For a given observed value y , the z -score may be determined using Eq. (1) below. One of the benefits of the z -score is the ability to compare dependent variables between groups when the independent variable is continuous (e.g., age) [19–23, 25, 26].

$$z = \frac{y - \text{mean}(\text{age})}{SD(\text{age})} \quad (1)$$

Step 5: Check the goodness-of-fit of the models

The z -score distribution should follow a normal distribution. If the data do not meet the model assumptions, then alternative power values for the polynomial model may be considered. Bootstrap replications were developed for computation.

Step 6: Calculate the RI

Lastly, the RI at various values of age is calculated based on the 5th and the 95th percentiles. Statistical analyses were performed using STATA 13.0 (Trial version) (Stata Corp. College Station, TX: Stata Press.) and MedCalc software (Trial version) (MedCalc Software Ltd. Ostend, Belgium).

Results

After outlier detection, this study included 4,051 and 2,667 patients for ALT and AST results, respectively. Among the ALT patients, 2,916 (72.0 %) were female, and the mean age $\pm$ SD was 35.08 ± 15.06 years. For AST, 1846 (69.2 %) were female, 821 (30.8 %) were male, and the mean age $\pm$ SD was 34.20 ± 14.95 years. The parameters had a non-normal distribution. Therefore, a logarithmic or a Box–Cox transformation was performed for each parameter. Then, the values were transformed to normal values. The Harris–Boyd method was used to check whether gender should be separated. The results indicated that ALT and AST measurements should be separated by gender.

Model selection with FP models

The FP models consist of different models according to the RA2 model selection procedures. The analysis procedure begins with an FP model of the maximum permitted complexity, determined by its dimension m_0 . The FP functions with $m_0=2$ (FP2) are sufficient for most applications when beginning the analysis [21]. We studied $m_0=2$ and $m_0=4$ to demonstrate the interpretation of complex model structures.

Table 1 presents the FP models, which start at $m_0=2$, for the AST values of females. First, the null models were compared with the FP2 models ($p=0.001$), indicating that age affects the female AST. Thus, the FP2 model was selected. Second, the FP2 model was compared with the linear model. The process was stopped because the result was not significant ($p=0.673$). Age had a linear relationship with AST values in females, and the linear model was selected as a fit model for calculating RIs. Table 2 summarizes the RIs for AST values of females for 10 years apart. Supplemental Data Table S1 presents the RIs for each age.

We found that $m_0=4$ (FP4) is the most complex function only in female ALT values in Table 3. The closed testing procedure begins by comparing the best fitting. The null model, linear model, and the best FP1, FP2, and FP3 models

Table 2: Reference intervals of AST values of female individuals calculated by FP models.

Age	Centiles of AST ^b	
	0.05 (90 % CI ^a)	0.95 (90 % CI ^a)
20	11 (11.6–12.1)	24 (23.7–24.7)
30	11 (11.7–12.1)	24 (24.3–25.1)
40	11 (11.8–12.2)	25 (24.8–25.6)
50	11 (11.8–12.4)	25 (25.1–26.3)
60	12 (11.8–12.6)	26 (25.3–27.0)
70	12 (11.7–12.8)	26 (25.6–27.8)
80	12 (11.7–13.0)	26 (25.8–28.6)
90	12 (11.6–13.3)	27 (26.0–29.5)

CI, confidence interval; SD, standard deviation. ^aBootstrap confidence interval (5,000 iterations). ^bCalculated by FP models: mean model= $16.975+0.025 \times \text{Age}$ and SD, model= $3.522 \times 0.012 \times \text{Age}$.

were compared against the more complex reference model FP4, at $\alpha=0.05$. First, the null model was compared with the best FP4 function, and it was significant. Thus, age was effective ($p<0.001$) on female ALT values. Then, the FP4 model was compared with the linear model. The test was significant, indicating that age had a non-linear relationship with female ALT values ($p<0.001$). Otherwise, the linear model would have been selected: compared with the best FP1 model, the FP4 model was selected because the tests were significant ($p<0.001$). Otherwise, the FP1 model should have been selected. A similar conclusion applies to the FP4 vs. the best FP2 comparison ($p=0.006$). Finally, FP4 was compared with the best FP3 model, indicating that FP4 was not significant ($p=0.098$). Thus, FP3 was selected as the best-fit model FP3.

According to the powers of the FP3 model (0, 0.5, 0.5), Table 4 summarizes the RIs for female ALT values for 10 years apart. Supplemental Data Table S2 provides the RIs for each age. Figure 1 illustrates the best-fitting model graphs.

According to the FP analysis results (for $m_0=2$), Table 5 presents separately the comparison between the null models and FP2 models for male ALT and AST values. The tests were not significant ($p=0.679$, $p=0.507$, respectively), and the

Table 3: FP models comparison for ALT values of females.

Model	df	Deviance	Deviance difference	p-Value ^a	Powers
Null	0	23,869.310	245.601	<0.001	
Linear	1	23,662.350	38.633	<0.001	1
m=1, FP1	2	23,653.200	29.490	<0.001	0.5
m=2, FP2	4	23,638.130	14.421	0.006	2 2
m=3, FP3	6	23,628.370	4.660	0.098	0 0.5 0.5
m=4, FP4	8	23,623.710	0.000	–	2 2 2 2

Table 1: FP models comparison for AST values of females.

Model	df	Deviance	Deviance difference	p-Value ^a	Powers
Null	0	15,089.260	18.201	0.001	
Linear	1	15,072.610	1.543	0.673	1
m=1, FP1	2	15,071.280	0.213	0.899	3
m=2, FP2	4	15,071.060	0.000	–	0.5 0.5

df, degrees of freedom and FP, fractional polynomial. ^ap-Values are comprising FP2 model, respectively.

df, degrees of freedom and FP, fractional polynomial. ^ap-Values are comprising FP4 model, respectively.

Table 4: Reference intervals of ALT values of female individuals calculated by FP models.

Age	Centiles of ALT ^b	
	0.05 (90 % CI ^a)	0.95 (90 % CI ^a)
20	5 (6.1–6.7)	21 (21.3–23.1)
30	6 (7.1–7.6)	24 (24.5–26.2)
40	6 (7.7–8.3)	26 (26.4–28.3)
50	7 (8.1–8.8)	27 (27.7–29.6)
60	7 (8.4–9.3)	28 (28.0–31.2)
70	7 (8.5–9.8)	29 (27.5–33.0)
80	8 (8.5–10.3)	30 (26.7–34.9)
90	8 (8.4–10.8)	30 (25.6–36.9)

CI, confidence interval; SD, standard deviation. ^aBootstrap confidence interval (5,000 iterations). ^bCalculated by FP models: mean model= $0.144+10.561\times\text{Log}(\text{Age})-0.183\times\text{Age}^{0.5}-0.183\times\text{Age}^{0.5}\times\ln(\text{Age})$ and SD, model= $1.561+2.600\times\text{Log}(\text{Age})+0.016\times\text{Age}^{0.5}+0.016\times\text{Age}^{0.5}\times\ln(\text{Age})$.

processes were stopped; the null models were selected. Therefore, age did not affect the male ALT and AST values. Therefore, RIs were calculated using traditional methods. The method is non-parametric because of the non-normality of values. Regardless of age, the lower and upper limits for ALT and AST were 8 (7–9) and 33 (32–33) and 12 (12–13) and 27 (27–28), respectively.

Discussion

We evaluated male and female ALT and AST levels to identify the best-fitting model and thus obtain RIs. Different methods were used to calculate RIs based on the relationship between age and liver function values. Female individuals

Table 5: FP models comparison for ALT and AST values of male individuals.

ALT					
Model	df	Deviance	Deviance difference	p-Value ^a	Powers
Null	0	9,543.251	2.320	0.679	
Linear	1	9,541.950	1.019	0.798	1
m=1, FP1	2	9,541.144	0.214	0.899	–1
m=2, FP2	4	9,540.931	0.000	–	0 2
AST					
Model	df	Deviance	Deviance difference	p-Value ^a	Powers
Null	0	6,871.416	3.328	0.507	
Linear	1	6,868.902	0.814	0.847	1
m=1, FP1	2	6,868.187	0.099	0.952	–2
m=2, FP2	4	6,868.088	0.000	–	–2 –2

df, degrees of freedom and FP, fractional polynomial. ^ap-Values are comprising FP4 model, respectively.

showed a non-linear relationship between age and ALT values; thus, FP models were used to calculate RIs for each age. A linear model was used for female AST values, whereas the traditional non-parametric method was used for male ALT and AST values. Our findings showed variations in different RI calculation methods.

The reference ranges of 18-year-old females' AST were 11–24 U/L. As age increases, the lower and upper limits of the range also increase. For the highest age of 92, the reference range was 12–27 U/L. Although ages with close gaps have slight changes, as the age gap increases, the difference widens, especially for the upper limit. The reference range for male AST is 12–27 U/L. Thus, men had higher values than women. For ALT, the reference ranges for women are 5–21 U/L and

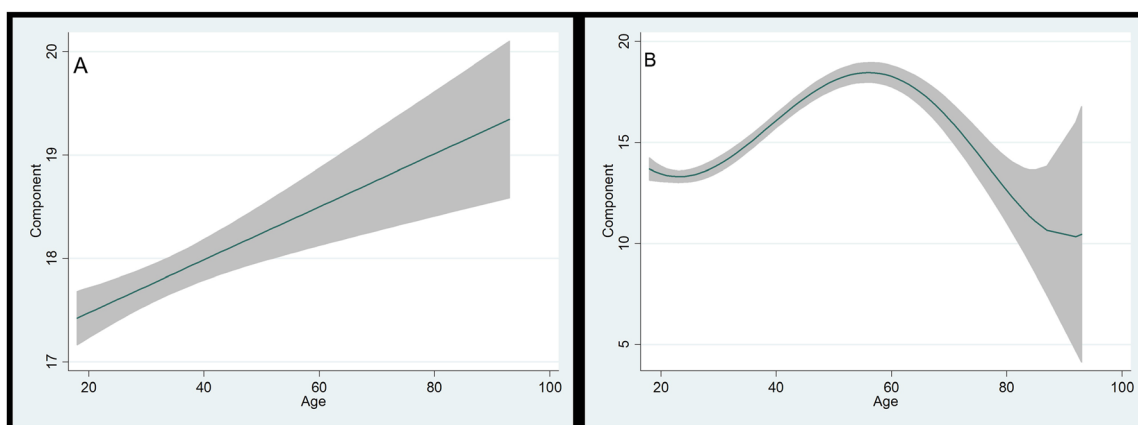


Figure 1: The best fitting model for AST and ALT of female individuals. This information is important for understanding the trends and patterns in AST and ALT values among females. One of the problems for the non-statistician is understanding the systematic approach used to compute the predicted mean and standard deviation, components of the Z score equation, which may vary as the independent variable changes over time (e.g., age) [18]. (A) AST tended to a linear increase in age and (B) ALT tended to a polynomial relation with age.

8–30 U/L for the minimum (18) and maximum (92) ages, respectively. According to age, the change in the lower and upper limits of the reference range was greater in ALT than in AST. For males, the range is 8–33 U/L. These values are within the accepted reference ranges. Therefore, age and sex-specific RIs are crucial in clinical practice to reduce the error rate of the post-analytical process, the correct grading of aminotransferase elevation, and evaluation of the fold increase or proportional values of abnormal aminotransferases [1, 3, 7–18]. By using these intervals, healthcare professionals can better assess whether a patient's test results fall within the expected range for their specific demographic group [9]. This assessment can aid in the diagnosis and monitoring of various conditions, such as hormonal imbalances or age-related changes in organ function [7, 8, 14].

Although the increase in aminotransferases guides the evaluation of liver diseases, such as chronic hepatitis C, autoimmune hepatitis, and nonalcoholic steatohepatitis, the exact limits on this issue are not completely clear [29]. Therefore, this study contributes to the literature significantly. Although different approaches are available to handle age in reference range calculations, researchers have traditionally divided RIs into age groups based on statistical significance and physiological relevance. Nevertheless, discrete RIs are insufficient to represent accurately the gradual change in analytes, such as ALT and AST, in response to age. The most common approach is to categorize age groups. However, to prevent information loss due to continuous variable categorization, FPs that keep the variable continuous and consider non-linearity have been recently proposed [17–23, 25, 26].

We compared the continuous and discrete RIs for only ALT and AST values for female because we calculated continuous RIs for them. We compare the results according to the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) values. Discrete RIs were calculated by Decision Tree (DT) method. The sub-groups according to the Decision Tree are used to calculate the reference intervals. The model with the smallest value for AIC and BIC is the best model for the data. For ALT female, a 39.5-year cut-off value was determined for discrete RIs and AIC and BIC values were 4,592.59 and 4,597.59, respectively. We calculated these values for FP model 4,577.94 and 4,587.94, respectively. According to the results, continuous RIs were better than discrete RIs. For AST female, a cut-off value could not be calculated by DT and this showed that continuous method was more precise.

Our study reveals the importance of calculating using the FP models and providing results for all ages in making successful and accurate interpretations. This type of RI was referred to as continuous RI. Many continuous RI studies have been conducted recently, particularly in the pediatric

population [7, 8, 14, 15], with several methodological studies conducted on modeling continuous RIs in addition to FP models [30–33]. Non-parametric models including quantile regression with smoothing splines, such as natural, restricted cubic, and B-splines, and semi-parametric models including lambda, mu, sigma and generalized additive models for location, scale, and shape are common methods [7, 8]. The use of FPs in the construction of RIs has increasingly become popular. Therefore, we selected FP models as a statistical method in this study. Chitty and Altman published the first study used in clinical science regarding size charts for fetal bones (radius, ulna, humerus, tibia, fibula, femur, and foot) after fitting FPs [27]. Kurmanavicius et al. used this method to produce ranges for cephalic index, head circumference, and occipitofrontal diameter [34]. They also used FPs to estimate the mean abdominal diameter, circumference, and femur length [35]. Zierk et al. calculated pediatric RIs for 15 biochemical analytes using FP models [13]. Our study is the first to calculate the continuous RIs for adults using FP models.

The validity and reliability of the calculated RIs were assessed using bootstrap resampling techniques, which allowed for the estimation of confidence intervals and evaluation of potential sources of bias. The results indicated that the RIs were robust and stable, with minimal variation observed across different resampling iterations. Therefore, the calculated intervals are reliable and can be confidently used for clinical decision-making. However, the calculation process may still have inherent limitations and sources of error, such as measurement errors or variations in the dataset used. These limitations should be considered when interpreting and applying the RIs in practice.

The strengths of our study are as follows. First, continuous RIs represent accurately variations in biomarker concentrations that occur with aging [7, 8, 13]. Second, our sample size is sufficient, though continuous RIs may require a larger sample size or subgroups to ensure statistical power and accuracy than partitioning RIs [7]. It is recommended that laboratories determine their own reference intervals (RIs) and regularly validate them. However, determining the RIs through traditional methods requires resources. For this reason, many laboratories use the reference intervals (RIs) recommended by manufacturers, but the values provided often do not align with population values [36, 37]. For this reason, considering the challenges in traditional RI calculations, the practicality of RI calculations made using hospital data is higher. In this study, the existing AST and ALT reference intervals used were calculated for each age, whereas this study provided more precise reference interval values based on age and sex variables.

Demographic variables such as age and sex require the calculation of reference ranges for the population as specific

as possible, reflecting the diversity among individuals. The traditional popRI model, in which population data is sorted from the lowest to the highest value to define RI, does not fully account for the natural biological variation of the measured entity [38]. However, it is emphasized that the genetic effects on RIs need to be investigated in more detail and that genotype-specific RIs should be produced [39]. Based on this information, we believe that the data obtained in this study will contribute to the ongoing transition towards individual reference ranges in clinical applications.

However, several limitations should be acknowledged. First, continuous RIs are not supported by laboratory information systems for clinical laboratory reporting due to mathematical functions. Furthermore, to guarantee accurate interpretation and use in clinical practice, the introduction of new RIs may require healthcare professionals complete training or education. Second, the process recommended by CLSI for determining reference intervals (RIs) is a direct method that involves the prior selection of the reference population [2]. However, due to the challenges encountered in the pre-analytical and analytical processes of this procedure, it is extremely difficult to carry it out properly. At the same time, it is time-consuming and also costly. For this reason, an indirect method is used that leverages routine laboratory data stored in the laboratory information system (LIS) to derive RIs [40, 41]. Therefore, the study was conducted in a single geographic region, which may limit the generalizability of the findings to other populations but the aim of this research is to investigate the applicability of fractional polynomial (FP) regression models as an alternative method for determining the reference range. Addressing these limitations and exploring potential sources of bias can enhance the validity and applicability of the research findings.

The study findings have important implications for clinical practice. By using age and sex-specific RIs, healthcare providers can make accurate diagnoses and treatment decisions, leading to improved patient outcomes. These RIs can also aid in the early detection of diseases or abnormalities, allowing for timely intervention and prevention of complications. Compared with discrete RIs, continuous RIs are more accurate [8]. Future research could explore the application of FP models in different populations, such as pediatric and geriatric patients, to determine whether the method performs equally well across different age groups. Additionally, incorporating additional covariates, such as BMI or medical history, into the regression analysis may further improve the accuracy of the RIs.

Research ethics: This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The local research Ethics Committee approved the study on 06/03/2024 (Decision number: 2024/251).

Informed consent: Not applicable.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Use of Large Language Models, AI and Machine Learning Tools: None declared.

Conflict of interest: The authors state no conflict of interest.

Research funding: None declared.

Data availability: The raw data can be obtained on request from the corresponding author.

References

1. Zierk J, Arzideh F, Haeckel R, Cario H, Frühwald MC, Groß HJ, et al. Pediatric reference intervals for alkaline phosphatase. *Clin Chem Lab Med* 2017;55:102–10.
2. Clinical and Laboratory Standards Institute (CLSI). Defining, establishing, and verifying reference intervals in the Clinical Laboratory; approved guideline, 3rd ed. Wayne (PA): Clinical and Laboratory Standards Institute; 2008. CLSI document EP28-A3c.
3. Colantonio DA, Kyriakopoulou L, Chan MK, Daly CH, Brinc D, Venner AA, et al. Closing the gaps in pediatric laboratory reference intervals: a CALIPER database of 40 biochemical markers in a healthy and multiethnic population of children. *Clin Chem* 2012;58:854–68.
4. Bildirici MA, Gülten S, Çalışgan NC. Determination of reference intervals of hemogram with advanced clinical parameters by indirect method on Sysmex XN-1000. *TJ Biochem* 2023;48:388–96.
5. Tewabe H, Amare T, Misganaw B. Establishment of reference interval for selected clinical chemistry parameters in apparently health adult in sekela and burie woredas, Ethiopia. *Clin Pract* 2021;18:1655–61.
6. Lala V, Zubair M, Minter DA. Liver function tests. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2023.
7. Hall A, Bohn MK, Wilson S, Higgins V, Adeli K. Continuous reference intervals for 19 endocrine, fertility, and immunochemical markers in the CALIPER cohort of healthy children and adolescents. *Clin Biochem* 2021;94:35–41.
8. Yan R, Peng Y, Hu L, Zhang W, Li Q, Wang Y, et al. Continuous reference intervals for 21 biochemical and hematological analytes in healthy Chinese children and adolescents: the PRINCE study. *Clin Biochem* 2022;102:9–18.
9. Temel GO, Kanik EA, Yıldırım Derici D, Ovla HD, Kalafat H. Determination of sub-groups in reference interval calculations and an application. *TJ Biochem* 2015;40:197–204.
10. Zhang GM, Guo XX, Zhu BL, Zhang GM, Bai SM, Wang HJ, et al. Establishing reference intervals of aspartate aminotransferase-to-platelet ratio index for apparently healthy elderly. *Clin Lab* 2016;62:135–40.
11. Guidi GC, Salvagno GL. Reference intervals as a tool for total quality management. *Biochem Med* 2010;20:165–72.
12. Wilson SM, Bohn MK, Madsen A, Hundhausen T, Adeli K. LMS-based continuous reference percentiles for 14 laboratory parameters in the CALIPER cohort of healthy children and adolescents. *Clin Chem Lab Med* 2023;61:1105–15.
13. Zierk J, Baum H, Bertram A, Boeker M, Buchwald A, Cario H, et al. High-resolution pediatric reference intervals for 15 biochemical analytes described using fractional polynomials. *Clin Chem Lab Med* 2021;59:1267–78.
14. Asgari S, Higgins V, McCudden C, Adeli K. Continuous reference intervals for 38 biochemical markers in healthy children and

- adolescents: comparisons to traditionally partitioned reference intervals. *Clin Biochem* 2019;73:82–9.
15. Zierk J, Arzideh F, Rechenauer T, Haeckel R, Rascher W, Metzler M, et al. Age- and sex-specific dynamics in 22 hematologic and biochemical analytes from birth to adolescence. *Clin Chem* 2015;61:964–73.
 16. Li K, Hu L, Peng Y, Yan R, Li Q, Peng X, et al. Comparison of four algorithms on establishing continuous reference intervals for pediatric analytes with age-dependent trend. *BMC Med Res Methodol* 2020;20:1–14.
 17. Regier M, Parker R. Smoothing using fractional polynomials: an alternative to polynomials and splines in applied research: smoothing using fractional polynomials. *WIREs Comput Stat* 2015;7:275–83.
 18. DeVore GR, Epstein A. Computing Z-score equations for clinical use to measure fetal umbilical vein size and flow using six independent variables of age and size. *J Ultrasound Med* 2022;41:1949–60.
 19. Zhang Z. Multivariable fractional polynomial method for regression model. *Ann Transl Med* 2016;4:174.
 20. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol* 1999;28:964–74.
 21. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. *J R Stat Soc Series C: Appl Stat* 1994;43:429–67.
 22. Royston P, Sauerbrei W. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. *J Roy Stat Soc* 1999;162:71–94.
 23. Royston P, Sauerbrei W. Fractional polynomials for one variable, MFP: multivariable model-building with fractional polynomials. In: Balding DJ, Cressie NAC, Fitzmaurice GM, Johnstone IM, Molenberghs G, Scott DW, et al, editors. *Multivariable Model-Building*, 1st ed. New York: John Wiley & Sons; 2008:2–4 pp.
 24. Dinç H, Ata Tutkun N. Cox regression model with fractional polynomials: an application on prostate cancer data set. *Türkiye Klinikleri J Biostat* 2020;12:38–52.
 25. Sauerbrei W, Royston P, Meier-Hirmer C, Benner A. Multivariable regression model building by using fractional polynomials: description of SAS, STATA, R programs. *Comput Stat Data Anal* 2006;50:3464–85.
 26. Ambler G, Royston P. Fractional polynomial model selection procedures: investigation of type I error rate. *J Stat Comput Simulat* 2001;69:89–108.
 27. Altman DG, Chitty LS. Design and analysis of studies to derive charts of fetal size. *Ultrasound Obstet Gynecol* 1993;3:378–84.
 28. Royston P, Wright EM. A method for estimating age-specific reference intervals (“normal ranges”) based on fractional polynomials and exponential transformation. *J Roy Stat Soc A* 1998;161:79–101.
 29. Lee TH, Kim WR, Poterucha JJ. Evaluation of elevated liver enzymes. *Clin Liver Dis* 2012;16:183–98.
 30. Zierk J, Arzideh F, Haeckel R, Rascher W, Rauh M, Metzler M. Indirect determination of pediatric blood count reference intervals. *Clin Chem Lab Med* 2013;51:863–72.
 31. Wan X, Qu Y, Huang Y, Zhang X, Song H, Jiang H. Nonparametric estimation on age-specific reference percentile curves with radial smoothing. *Contemp Clin Trials* 2012;33:13–22.
 32. Borghi E, Onis M, Garza C, Van Den Broeck J, Frongillo EA, Grummer-Strawn L, et al. Construction of the World Health Organization child growth standards: selection of methods for attained growth curves. *Stat Med* 2006;25:247–65.
 33. Bussler S, Vogel M, Pietzner D, Harms K, Buzek T, Penke MN, et al. New pediatric percentiles of liver enzyme serum levels (alanine aminotransferase, aspartate aminotransferase, gamma-glutamyltransferase): effects of age, sex, body mass index, and pubertal stage. *Hepatology* 2018;68:1319–30.
 34. Kurmanavicius J, Wright EM, Royston P, Wissner J, Huch R, Huch A, et al. Fetal ultrasound biometry: 1. Head reference values. *Br J Obstet Gynaecol* 1999;106:126–35.
 35. Kurmanavicius J, Wright EM, Royston P, Zimmermann R, Huch R, Huch A, et al. Fetal ultrasound biometry: 2. Abdomen and femur length reference values. *Br J Obstet Gynaecol* 1999;106:136–43.
 36. Örkmez M, Tarakcioglu M. Determination of reference intervals of Biochemistry parameters in healthy individuals in gaziantep province. *Eur J Ther* 2023;29:173–8.
 37. Bakan E, Polat H, Ozarda Y, Ozturk N, Baygutaalp NK, Umudum FZ, et al. A reference interval study for common biochemical analytes in Eastern Turkey: a comparison of a reference population with laboratory data mining. *Biochem Med* 2016;26:210–23.
 38. Coşkun A, Sandberg S, Unsal I, Topcu DI, Aarsand AK. Reference intervals revisited: a novel model for population-based reference intervals, using a small sample size and biological variation data. *Clin Chem* 2024:hvae109.
 39. Ozarda Y. Reference intervals: current status, recent developments and future considerations. *Biochem Med* 2016;26:5–16.
 40. Ozarda Y, Ichihara K, Barth JH, Klee G. On behalf of the CRIDL on Reference intervals and Decision Limits, IFCC. Protocol and standard operating procedures for common use in worldwide multicenter study on reference values. *Clin Chem Lab Med* 2013;51:1027–40.
 41. Jones GRD, Haeckel R, Loh TP, Sikaris K, Streichert T, Katayev A, et al. IFCC Committee on Reference Intervals and Decision Limits. Indirect methods for reference interval determination – review and recommendations. *Clin Chem Lab Med* 2018;57:20–9.

Supplementary Material: This article contains supplementary material (<https://doi.org/10.1515/tjb-2024-0086>).