Opinion Paper

Gorkem Sezgin*, Tze Ping Loh and Corey Markus

Functional reference limits: a case study of serum ferritin

<https://doi.org/10.1515/labmed-2020-0127>Received October 28, 2020; accepted February 22, 2021;
published online March 5, 2021

Abstract: Reference intervals depend on the distribution of results within a reference population and can be influenced by subclinical disease. Functional reference limits present an opportunity to derive clinically relevant reference limits from routinely collected data sources, which consist of mixed populations of unhealthy and healthy groups. Serum ferritin is a good example of the utility of functional reference limits. Several studies have identified clinically relevant reference limits through examining the relationship between serum ferritin and erythrocyte parameters. These ferritin functional limits often represent the inflection point at which erythrocyte parameters change significantly. Comparison of ferritin functional reference limits with those based on population distributional reference limits reveals that the lower reference limit may fall below the point at which patients become clinically unwell. Functional reference limits may be considered for any biomarker that exhibits a correlated relationship with other biomarkers.

Keywords: functional reference limit; harmonization; reference interval; serum ferritin.

Introduction

Reference intervals are often a representation of the ‘non-diseased’ population and reported as the 2.5th and 97.5th percentiles [1, 2]. Sampling a non-diseased population can be challenging, as any sampling approach to derive reference intervals will inherently have unhealthy subjects

within the population [3]. Although various study designs and statistical approaches can be utilized to alleviate such bias being reflecting into reference intervals, populations with high incidence of subclinical disease are particularly challenging. Additionally, clinical conditions that exist in a wide spectrum of severity, from physiological to pathological, makes it hard to clearly delineate the subpopulations to be excluded.

Serum ferritin is often reported as having lower reference intervals in females compared to males [4]. However, harmonization studies reporting such differences suggest the possibility of subclinical diseases influencing the reference intervals [5]. As shown in Table 1, serum ferritin reference intervals from various studies are persistently lower for females. This notion was the focus of a review by Rushton and Barth [4], who suggested that the difference between genders for the serum ferritin reference intervals is due to the high incidence of subclinical iron deficiency among the general adult female population. The extent of any subclinical disease in a population is generally unknown and its impact therefore difficult to estimate. Iron deficiency is a global health concern and is highly prevalent among the general female population [6] and difficult to identify [7].

Often, differences between sexes are conditional on age: certain age groups may show larger differences compared to others. Serum ferritin reference intervals often present with differences between sexes and age, with lower levels among females during adolescence and early adulthood, compared to childhood and older ages [8]. One of the foremost theories is the effect of onset of menses resulting in the differences observed. Opinion varies regarding levels of serum ferritin observed during the stages of the menstrual cycle. A study by Kim et al. [9] showed different levels in serum ferritin biomarkers during different stages of the menstrual cycle. In contrast, while the study by Belza et al. [10] had fewer subjects, their study design included longitudinal sampling of subjects and suggested a single blood sample at any time in the cycle would identify iron deficiency. How hormonal changes with menstruation influences serum ferritin reference intervals is not clear. It is possible that the effect is negligible, and the reported reference intervals, being lower

*Corresponding author: Gorkem Sezgin, Faculty of Medicine, Health, and Human Sciences, Macquarie University, Sydney, Australia, Phone: +61 298502490, E-mail: gorkem.sezgin@mq.edu.au

Tze Ping Loh, Department of Laboratory Medicine, National University Hospital, Buona Vista, Singapore

Corey Markus, International Centre for Point-of-Care Testing, Flinders University, Adelaide, Australia. <https://orcid.org/0000-0002-5594-9737>

Table 1: Select reference interval studies from 2020 reporting serum ferritin results by direct and indirect approaches.

Reference	Region	Gender/age	Lower limit	Upper limit	Design (no. of samples)
Topić and Krleža (2020) [12]	Croatia	1–<16 y	5.8, µg/L	37.4, µg/L	Non-diseased, direct, distribution based (n=295)
		Male 16–<18 y	25.6, µg/L	88.0, µg/L	
		Female 16–<18 y	4.8, µg/L	42.5, µg/L	
Yeung et al. (2020) [13]	Hong Kong	Male >12 y	34, µg/L	271, µg/L	Non-diseased, indirect, distribution based (n=160)
		Female >12 y	6, µg/L	126, µg/L	
Wang et al. (2020) [14]	China	Parametric			Indirect, distribution based (n=14,310)
		Male	66.1, µg/L	562, µg/L	
		All females	3.6, µg/L	270, µg/L	
		Female <50 y	3.3, µg/L	148, µg/L	
		Female ≥50 y	17.3, µg/L	303, µg/L	
		Non-parametric			
		Male	65.0, µg/L	571, µg/L	
		All females	4.0, µg/L	254, µg/L	
		Female <50 y	4.0, µg/L	152, µg/L	
		Female ≥50 y	16.0, µg/L	304, µg/L	
Borai et al. (2020) [15]	Saudi Arabia	Male	6, µg/L	296, µg/L	Non-diseased, direct, distribution based (n=826)
		Female 18–44 y	3.6, µg/L	102, µg/L	
		Female 45–65 y	1.6, µg/L	163, µg/L	
Omuse et al. (2020) [16]	Kenya	18–65 y	8.8, µmol/L	28.9, µmol/L	Non-diseased, direct, distribution based (n=533)

for adolescent and adult women in reproductive ages is an accurate reflection of physiological changes. As shown in Figure 1, percentiles among females begin to decline at ages corresponding to the onset of menses. It is also possible that women have a greater prevalence of sub-clinical iron deficiency due to not replenishing iron lost during the menstrual cycle or previous childbirth, which will inevitably reflect on the percentiles calculated as reference intervals [4]. Another potential reason for the age and gender difference of serum ferritin levels could be differences in iron absorption, although more studies

comparing iron absorption differences among adults are needed [11].

Functional reference limits

Statistical analysis is limited to the available information, and although a few numbers of outliers can be addressed, inclusion of a high number of pathological results is difficult to reliably address. Statistical analysis based on population distributions could therefore be vulnerable to

Figure 1: Dynamics of ferritin from birth into early adulthood for both genders (18,770 females; 24,445 males).

Results of all ferritin results from an 11-year data extraction from a pediatric hospital laboratory data base, for patients aged 0–18 years. Symbol colors are logarithmically proportional to the number of ferritin results; as indicated by the color chart to the right. Dashed lines are the ferritin reference interval; 20–200 µg/L for 3-months to 18-years of age in both genders. Solid lines show the 25th, 50th (median) and 75th percentiles.

biases within the population distribution and could result in reference intervals that are not clinically “normal”, even though it may be statistically normally distributed. Additionally, transformations to achieve normally distributed data need to be carefully performed to avoid further inadvertent inclusion of pathological results.

An emerging approach to deriving reference intervals is by utilizing both diseased and non-diseased populations to determine “functional” reference limits. Here, the word “functional” can describe both biological and statistical relationships. Physiological processes occur as functional relationships, with changes in one component affecting another related component(s). Homeostasis is a physiological state where the interactions between biomarkers occur within normal operating parameters; with disease, these physiological interactions become pathological and the components of physiological processes show correlated disturbances. Functional reference limits can be defined as the numerical threshold at which the relationship between physiologically related biomarkers changes significantly. Physiologically, it may represent a functional change in physiologic state (e.g., failure of homeostasis) or pathology. Statistically, this is often represented by a change in curvature of the function curve. Often, the functional relationship is specific for a particular clinical condition/physiology. Functional reference limits are calculated by correlating two interrelated biomarkers with each other, displayed visually in a graph, and mathematically modeling the correlation of the biomarkers at the point where the relationship becomes non-linear as displayed in Figure 2.

Functional reference limits are particularly suitable for serum ferritin, with clinical relevance to iron deficiency anemia. The earliest ferritin reference intervals were derived from ferritin levels at which an absence of stainable iron stores in bone marrow aspirate (considered the gold standard) are observed, which could be regarded as the earliest example of a “functional reference limit” [17]. Given the difficulty of sample collection and subjectivity of stainable bone marrow iron stores, a functional ferritin reference limit derived from modern ferritin assays against well characterized and automated measurement of erythrocyte parameters, easily extracted from laboratory databases, and able to account for confounding influences described previously, would provide a sound evidence base for a functional ferritin threshold for iron deficiency.

A decrease in serum ferritin is followed similarly by decreases in erythrocyte biomarkers, as per the progression of iron deficiency into anemia [7]. In treatment of iron deficiency anemia, hemoglobin increases with increasing serum ferritin, up to a serum ferritin threshold above which the hemoglobin concentration “plateaus”. This threshold or

Figure 2: Example of a correlation graph showing the association between two related biomarkers, alongside their smoothed histograms showing the distribution of results.

The example is based on a simulated functional relationship between biomarker A and B. Biomarker A is based on a log-normal distribution, while Biomarker B is a non-linear transformation of biomarker A, with added noise. Two variations of biomarker A are simulated, one with 5% of the sample ($n=10,000$) having subclinical disease (orange), and the other unaffected/normal (gray). As biomarker A results increase, biomarker B increases until a point where it starts to plateau and this point is considered the functional threshold. The functional threshold is demonstrated by the dashed-magenta line, and the 2.5th percentile is demonstrated by the dashed-black line. As shown by the Figure, the inclusion of a 5% diseased sample has pulled the 2.5th percentile downwards, whilst the functional threshold remains unaffected.

pseudo-plateau point can be described as the functional reference limit, as shown in Figure 3. Åsberg et al. [18] described the functional correlation between hemoglobin and serum ferritin for ambulant patients, finding an abrupt downward shift in hemoglobin concentrations when serum ferritin reduced below 20 µg/L. Abdullah et al. [19] assessed the functional correlation between serum ferritin and hemoglobin for children aged between 12 and 36 months, reporting a serum ferritin value of <17.9 µg/L is when functional iron deficiency may occur, based on the deteriorating levels in hemoglobin. Markus et al. [20] correlated serum ferritin with several erythrocyte parameters: hemoglobin, mean corpuscular volume, mean corpuscular hemoglobin concentration, mean corpuscular hemoglobin, and red cell distribution width. From their observations, they proposed 26 µg/L for children between 4 months and <13 years, and at least 39 µg/L for adolescents as lower functional limits. Foy et al. [21] used adult serum ferritin and erythrocyte results from hospitals, reporting an approximately five percent decline in erythrocyte indices between 10 to 25 ng/mL serum

ferritin. Similarly, Sezgin et al. [22] utilized pediatric serum ferritin results and correlated erythrocyte results to calculate a functional reference limit for serum ferritin, finding a “hematological plateau” occurring near 10 $\mu\text{g/L}$ (Figure 3).

Before computational resources were as advanced as today, Agha et al. [23] illustrated the correlation between transferrin saturation and serum ferritin as a scatterplot. Although they did not define it as a functional reference limit, the authors found serum ferritin levels below 10 ng/mL is associated with a transferrin saturation less than 15%. Åsberg et al. [18] had a much larger dataset, and used quantile regression modeling to calculate 10th, 50th, and 90th percentiles for the relationship between hemoglobin and serum ferritin, in addition to multivariable regression spline to determine the simplest non-linear transformation for serum ferritin. These more advanced statistical processes are possible through the improved computational power today.

Abdullah et al. [19] used restricted cubic spline regression to determine the association between serum ferritin and hemoglobin, in their case, also adjusting for potential confounders. Utilizing the spline model, they defined the point at which the relationship plateaus as: the serum ferritin level which predicts the maximum increase in hemoglobin as a function of serum ferritin. Additionally, they correlated the hemoglobin cut-off value corresponding to anemia with serum ferritin.

Figure 3: Correlation between serum ferritin ($\mu\text{g/L}$) and hemoglobin (g/L) results for a female population between 13 and 18 years, used to calculate functional reference limits, from Sezgin et al. [22].

Scatterplot presents the raw test result values, while the solid-line going through the scatterplot is the fitted quantile regression model. The lighter line closely following the quantile regression model is a median spline fitted to show the distribution of the scatterplot. The threshold at which the pseudo-plateau begins to occur is defined as the functional reference limit, which was designated as 10 $\mu\text{g/L}$ serum ferritin. Vertical lines at 10, 15, and 20 $\mu\text{g/L}$ serum ferritin are provided for reference. Reprinted with permission.

Markus et al. [20] calculated a non-linear quantile regression to the median with three-parameter exponential equations to the erythrocyte parameters, as a function of ferritin. They derived their functional limits through determining the point at which the modeled hematological values reached a near asymptote (the point at which the pseudo-plateau occurred, in their case). Foy et al. [21] utilized a cubic smoothing spline modeling, with their analysis stratified by age and gender. Their limits were defined as the serum ferritin level at which a greater than 5 or 10% decline occurred for erythrocyte parameters; with the exception of hemoglobin distribution width and red cell volume, for which an increase in these parameters relative to the population mean in stable ferritin region (defined as 50–150 ng/mL) were designated as the limits. Sezgin et al. [22] calculated a fractional polynomial regression model to the median. Their functional reference limit was defined as a plateau occurring when the erythrocyte value difference between subsequent correlated ferritin values becomes less than one.

Although the general principle remains the same, these studies use different approaches to determine the functional limits (Table 2). The first difference arises on the modeling of the results, and the second difference arises from the definition of the “limit” at which the functional decline or increase occurs. In determining functional reference limits for any biomarker, the first aim is to determine the shape of the association between the biomarkers, ideally through exploratory data analysis. In Figure 3, the scatterplot suggests a singular inflection point. Further exploratory modeling through median spline suggests that after this inflection points, a plateau-like shape occurs. Based on this information, a suitable model to describe this association is selected, in which case a fractional polynomial regression model was deemed appropriate. In contrast, Abdullah et al. observed four to five knots rather than a singular inflection points, with varying levels of changes occurring within the pseudo-plateau [19]. In such case, a restricted cubic spline model was deemed appropriate to model the association between serum ferritin and hemoglobin. Once the model is formed, it is used to determine the “inflection point”, representing a point or region where a significant change occurs, which may be of statistical significance (e.g. plateau), and/or clinical significance (e.g., significant decline in hemoglobin in relation to changing serum ferritin values). The inflection point may be defined through various approaches, as listed in Table 2. Åsberg et al. and Foy et al. utilized a statistical approach (statistical significant difference between groups, and >5–10% decline in erythrocyte values, respectively) while Abdullah et al. formed an additional clinical criteria,

defining the inflection point as the serum ferritin value corresponding to a Hb value of 110 g/L [18, 19, 21]. Currently, there is no gold standard on the best approach to defining an inflection point, although the different approaches seem to produce comparable results.

Conceivably, functional reference limits can be determined for biomarkers beyond serum ferritin, including calcium homeostasis, and other micronutrient deficiency/excesses. Vitamin D deficiency is also common among the general population. There is no clear consensus for the cutoffs to define vitamin D deficiency, and deficiencies present at varying levels by age and gender [24–26]. Functional reference limits can provide clinically relevant values through modeling the association between vitamin D, parathyroid hormone, and calcium [25, 26]: all of which play an important role in calcium and phosphate homeostasis, with implications for development of calcium disturbances, and osteoporosis [27]. Folate levels can vary considerably between different populations as a result of different dietary intakes [28, 29]. Through the association between folate and homocysteine, a robust functional reference limit which is not influenced by dietary factors can be estimated [30].

Clinical use of functional reference limits

The main utility of functional reference limits lies as a means of supporting evidence for reference intervals, and as potential clinical decision limits for specific clinical conditions. In the case of serum ferritin, functional limits provide valuable information on the effect of inclusion of

subclinical iron deficiency on female reference intervals, as well as providing a clinically relevant limit for iron deficiency anemia. Sezgin et al. correlated serum ferritin with various erythrocyte biomarkers, and determined a functional reference limit for each [22]. At the higher end point, a serum ferritin level near 10–11 µg/L was associated with a sharp decline in hemoglobin levels, which is suggestive of incipient iron deficiency anemia. At the lower end point, a serum ferritin level near 6–8 µg/L was correlated with declining mean corpuscular volume, which is indicative of microcytic anemia. Clinically, such information may provide insight into physiological/homeostatic limits and severity of suspected iron deficiency. However, functional reference limits do not indicate when a patient becomes iron deficient, which would require either reference intervals or clinical decision limits.

An interesting observation when comparing serum ferritin functional reference limits to the distribution-based reference intervals (calculated by laboratories and harmonization studies) is that the functional limits are higher compared to the lower reference limit for serum ferritin (Tables 1 and 2), and is more evident for females. Functional limits calculated for serum ferritin are determined by associations with erythrocyte parameters: the values derived are statistically relevant thresholds that represent the iron stores at which erythropoiesis starts to be impacted. As such, these functional limits are expected to be higher than population distribution derived reference intervals.

It has been thought previously that reference intervals for females were too low [4], which may be an artifact resulting from the high prevalence of (subclinical) iron deficiency among the general female population [31]. When lower reference intervals for serum ferritin are used for

Table 2: Various models and definitions used to calculate functional reference limits for serum ferritin.

Reference	Functional lower reference limit	Model	Definition(s)
Åsberg et al. (2013) [18]	Female <20, µg/L Male <30, µg/L	Quantile regression (n=12,270)	Statistical significant difference (p<0.05) between various groups based on serum ferritin
Abdullah et al. (2017) [19]	Children 12–36 months: Hb plateau point: 17.9–23.7, µg/L Hb 110 g/L: 2.4–4.6, µg/L	Restricted cubic spline (n=1,257)	Hb plateau point; serum ferritin corresponding to a Hb of 110 g/L
Markus et al. (2019) [20]	4 months – <13 y: 24, µg/L 13–19 y: 39, µg/L	Quantile regression (n=54,896)	Hb plateau point
Foy et al. (2020) [21]	Age-averaged male ~35, ng/mL Female 18–44 y: ~15, ng/mL Female 45–65 y: ~25, ng/mL	Smoothed curve (n=58,451)	>5–10% decline for most erythrocyte indices; increase in measure relative to population mean in stable ferritin region (for red cell distribution width and hemoglobin distribution width).

assessing whether a patient may have iron deficiency, patients may be erroneously misclassified as being iron replete, when in fact they may already have iron deficiency. Although cut-off values for serum ferritin coincide better with the functional limits [8] and are claimed to be used instead of reference intervals for diagnosing iron deficiency [32] the question of the accuracy of lower serum ferritin reference interval remains.

Differential diagnosis of iron deficiency based on clinical symptoms is difficult, and serum ferritin is the more useful indicator of this condition [7]. Insufficient iron levels are particularly debilitating for young children even before the onset of anemia, with deficiencies thought to be associated with cognitive developmental impairment [33]. If distribution based percentiles are too low, as is suspected among young females, then iron deficiency may be underdiagnosed in the clinic. Sezgin et al. [22] compared distribution based reference intervals to functional reference limits using the same pediatric population. Their results show that percentiles calculated from the same population can fall below the functional limit indicative of iron deficiency erythropoiesis. Similarly, comparing their reference interval study to functional reference limit study, Parkin et al. reached to the conclusion that the proportion of misclassification due to using lower reference interval compared to functional reference limit for iron deficiency may be higher [5].

Functional reference limits and clinical decision limits

Functional reference limits and clinical decision limits share many similarities in their utility: they can both be used to determine whether a laboratory test result is indicative of a disease or deteriorating condition. Likewise, upper limits and lower limits for a biomarker are determined independently, unlike reference intervals in which the two limits are related to the distribution of results within the population it is being derived from.

The main difference between functional and clinical limits lies in their approach to deriving the limit. Clinical decision limits distinguish diseased population from non-diseased population and arbitrarily determine an 'acceptable' point or threshold for disease (often based on sensitivity and specificity) [34]. Functional reference limits on the other hand use a pool of diseased and non-diseased population to statistically determine the point when homeostasis is disturbed. Functional reference limits therefore do not necessarily have to have an

association with disease: they can be used as an indicator of deteriorating condition.

In the case of serum ferritin, a clinical decision limit may indicate when iron deficiency should be suspected, is moderate, or severe. Functional reference limits, as correlated by erythrocyte parameters, indicates at what level of serum ferritin the patient may become anemic, resulting from iron deficiency.

Limitations of functional reference limits

Functional reference limits require biomarkers to have biological associations with others such that change in one results in a change, or a cascade of changes in others. Many hematological biomarkers exist in a homeostatic state with one another, often showing correlated relationships. Increase in cellular iron levels for example is associated with increased ferritin concentrations, and vice versa [35]. However, not all biomarkers have such associations and correlation with others, and mostly act without affecting other biomarkers. For example, prostate specific antigens are currently not known to have any associations with other biomarkers that is indicative of a specific condition or disease. Functional reference limits can therefore not be determined in such cases.

Unlike reference intervals, functional reference limits are independent of one another, such that a lower functional reference limit and an upper functional limit are calculated separately. In some cases, the correlated biomarkers used for calculating lower functional reference limits cannot be used to calculate upper functional limits and vice versa. For example, a lower functional reference limit for serum ferritin can be calculated through its association with hemoglobin during iron deficiency anemia. On the upper end, it is not clear if hemoglobin can be reliably correlated with serum ferritin. Initial investigations suggest a relationship may be present due to anemia of chronic disease, with an inverted-U shaped association. Researches should therefore ensure that selected biomarkers have a homeostatic relationship with one another.

Sample size considerations also need to be taken into account when calculating functional reference limits. The number of samples required for accurate modeling is often higher than would be required for a reference intervals calculated from non-diseased populations, as a mixed sample of healthy and unhealthy populations are required. A further challenge (also faced by reference interval

calculation) is the requirement for partitioning, often for age and sex, which demands a larger overall number of samples. Not accounting for such differences could potentially lead to misclassification and poorer patient outcomes [36]. Often, different ages are grouped together based on the similarity of values between the ages [37, 38]. However, age grouping can potentially mask normal physiological trends and result in inaccurate reference intervals, particularly if the age groups are too broad [39]. Functional reference limits can be calculated through readily available information in many laboratory information systems, which often contain a large sample of results. Researchers calculating functional reference limits should however ensure the sample sizes are appropriate for an accurate model to be fitted, particularly if age- and/or sex-partitioning is required.

Special consideration: pregnancy

Pregnancy is a time of numerous physiological changes and daily iron requirements can increase from 1 to 6 mg/day with advancing gestation [40]. Iron deficiency with or without anemia is particularly prevalent in pregnant women, with a recent Australian study finding at first trimester presentations to the antenatal clinic, 31.3% and 5.7% having iron deficiency or iron deficiency anemia respectively [41].

A recent Australian study by Randall et al. [42] examined the relationship between maternal hemoglobin levels ≤ 20 weeks of gestation and pregnancy outcomes. Their study confirmed earlier work on a U-shaped relationship, with those women which had low or high hemoglobin having a greater risk of adverse outcomes at delivery than those with a normal hemoglobin [43]. Recommendations for routine iron supplementation in pregnancy vary and can depend on the local prevalence of iron deficiency [40, 44]. Crispin et al. [45] found serum ferritin measured in the first trimester was predictive of anemia in the pre-delivery period and can identify women for iron therapy and may be better than relying on hemoglobin alone.

Daru et al. [46] has thoroughly reviewed ferritin thresholds for definition of iron deficiency in pregnancy, and found cutoffs ranged from <6 to <60 $\mu\text{g/L}$. Daru et al. also reiterates the two commonly used cutoffs, <12 and <15 $\mu\text{g/L}$ were introduced into international guidelines by consensus agreements back in 1998 and 2001 respectively [46–48]. Daru et al. aimed criticism at the ferritin cutoff <30 $\mu\text{g/L}$ which has been incorporated into national guidelines determined from the 1998 study of van den Broek et al. [49] in which 47% of recruited

participants had HIV, and the small sample size limited the estimation the effect of HIV status on serum ferritin levels [46].

Serum ferritin may be predicative of pre-delivery anemia [42] and disturbances in hemoglobin levels are associated with poorer pregnancy outcomes [45]. A strong evidence based ferritin threshold for iron deficiency in pregnancy is urgently required: current ferritin cutoffs are derived from historical studies measured by assays likely no longer in use [46] against the backdrop of varying recommendations, both for and against, the use of routine iron supplementation [40, 44].

Summary and outlook

Functional reference limits provide a conceptual framework and flexible modeling approach by utilizing both healthy and unhealthy samples whilst providing clinically relevant reference limits. For serum ferritin, lower functional limits determined through its correlation with erythrocyte biomarkers represent changes which occur during iron deficiency anemia. Functional reference limits can be calculated for any biomarker which presents with a statistical relationship with other correlated biomarker(s).

Research funding: None declared.

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

Competing interests: Authors state no conflict of interest.

References

1. PetitClerc C, Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 2. Selection of individuals for the production of reference values. Clin Chim Acta 1987;170:S1–11.
2. Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. Clin Chim Acta 1987;170:S13–32.
3. Horn PS, Feng L, Li Y, Pesce AJ. Effect of outliers and nonhealthy individuals on reference interval estimation. Clin Chem 2001;47: 2137–45.
4. Rushton DH, Barth JH. What is the evidence for gender differences in ferritin and haemoglobin? Crit Rev Oncol Hematol 2010;73:1–9.
5. Parkin PC, Hamid J, Borkhoff CM, Abdullah K, Atenafu EG, Birken CS, et al. Laboratory reference intervals in the assessment of iron status in young children. BMJ Paediatr Open 2017;1:e000074.
6. Coad J, Conlon C. Iron deficiency in women: assessment, causes and consequences. Curr Opin Clin Nutr Metab Care 2011;14: 625–34.

7. Soppi ET. Iron deficiency without anemia – a clinical challenge. *Clin Case Rep* 2018;6:1082–6.
8. Worwood M. Indicators of the iron status of populations: ferritin. [Internet]. Available from: https://www.who.int/nutrition/publications/micronutrients/anaemia_iron_deficiency/9789241596107_annex2.pdf?ua=1 [Accessed 10 Feb 2020].
9. Kim I, Yetley EA, Calvo MS. Variations in iron-status measures during the menstrual cycle. *Am J Clin Nutr* 1993;58:705–9.
10. Belza A, Henriksen M, Ersbøll AK, Thilsted SH, Tetens I. Day-to-day variation in iron-status measures in young iron-deplete women. *Br J Nutr* 2005;94:551–6.
11. Woodhead JC, Drulis JM, Nelson SE, Janghorbani M, Fomon SJ. Gender-related differences in iron absorption by preadolescent children. *Pediatr Res* 1991;29:435–9.
12. Zrinski Topić R, Leniček Krleža J. Verification of the Canadian laboratory initiative on paediatric reference intervals (CALIPER) reference values in Croatian children and adolescents. *Biochem Med* 2020;30:307–14.
13. Yeung DK-H, Chung JK-O, Ng SW-S, Pang EW-H, Tam DC-C. Validation of reference intervals of serum ferritin in Hong Kong. *J Hong Kong Inst Med Lab Sci* 2018;15:10.
14. Wang Q, Guo L, Lu Z, Gu J. Reference intervals established using indirect method for serum ferritin assayed on Abbott Architect i2000SR analyzer in Chinese adults. *J Clin Lab Anal* 2020;34:e23083.
15. Borai A, Ichihara K, Masaud A, Tamimi W, Bahijri S, Armbuster D, et al. Establishment of reference intervals for immunoassay analytes of adult population in Saudi Arabia. *Clin Chem Lab Med* 2020;58:1302–13.
16. Omuse G, Ichihara K, Maina D, Hoffman M, Kagotho E, Kanyua A, et al. Determination of reference intervals for common chemistry and immunoassay tests for Kenyan adults based on an internationally harmonized protocol and up-to-date statistical methods. *PLoS One* 2020;15:e0235234.
17. Krause JR, Stolc V. Serum ferritin and bone marrow biopsy iron stores. II. Correlation with low serum iron and Fe/TIBC ratio less than 15%. *Am J Clin Pathol* 1980;74:461–4.
18. Åsberg A, Mikkelsen G, Thorstensen K, Åsberg AE. Lower hemoglobin with lower ferritin: it is not just a question of anemia. *Scand J Clin Lab Invest* 2013;73:622–6.
19. Abdullah K, Birken CS, Maguire JL, Fehlings D, Hanley AJ, Thorpe KE, et al. Re-evaluation of serum ferritin cut-off values for the diagnosis of iron deficiency in children aged 12–36 months. *J Pediatr* 2017;188:287–90.
20. Markus C, Saxon B, Metz M. Ferritin as a functional biomarker of iron status in children and young adults. *Br J Haematol* 2019;184:640–2.
21. Foy BH, Li A, McClung JP, Ranganath R, Higgins JM. Data-driven physiologic thresholds for iron deficiency associated with hematologic decline. *Am J Hematol* 2020;95:302–9.
22. Sezgin G, Monagle P, Loh TP, Ignjatovic V, Hoq M, Pearce C, et al. Clinical thresholds for diagnosing iron deficiency: comparison of functional assessment of serum ferritin to population based centiles. *Sci Rep* 2020;10:18233.
23. Agha F, Akhter P, Khan RA. Serum ferritin levels in apparently healthy subjects. *J Pak Med Assoc* 1987;37:63–6.
24. Holick MF. Vitamin D deficiency. *N Engl J Med* 2007;357:266–81.
25. Atapattu N, Shaw N, Höglér W. Relationship between serum 25-hydroxyvitamin D and parathyroid hormone in the search for a biochemical definition of vitamin D deficiency in children. *Pediatr Res* 2013;74:552–6.
26. Rizzoli R, Boonen S, Brandi M-L, Bruyère O, Cooper C, Kanis JA, et al. Vitamin D supplementation in elderly or postmenopausal women: a 2013 update of the 2008 recommendations from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). *Curr Med Res Opin* 2013;29:305–13.
27. Nussey S, Whitehead S. The parathyroid glands and vitamin D [Internet]. *endocrinology: An integrated approach*. BIOS Scientific Publishers; 2001. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK24/> [Accessed 9 Oct 2020].
28. Galukande M, Jombwe J, Fualal J, Baingana R, Gakwaya A. Reference values for serum levels of folic acid and vitamin B12 in a young adult Ugandan population. *Afr Health Sci* 2011;11:240–3.
29. Schwettmann L, Berbu S. Reference interval and status for serum folate and serum vitamin B12 in a Norwegian population. *Clin Lab* 2015;61:1095–100.
30. Vos MJ, Joost van Pelt L, Kok MB, Dijck-Brouwer DAJ, Heiner-Fokkema MR, Dikkeschei LD, et al. Folate reference interval estimation in the Dutch general population. *Pract Lab Med* 2019;16:e00127.
31. Rushton DH, Dover R, Norris MJ, Gilkes JJH. Iron and hair loss in women; what is deficiency? This is the real question! *J Am Acad Dermatol* 2007;56:518–9.
32. Heath A-LM, Fairweather-Tait S, Worwood M. Reference limits for haemoglobin and ferritin. *BMJ* 2001;323:806.
33. Beard JL. Why iron deficiency is important in infant development. *J Nutr* 2008;138:2534–6.
34. Ozarda Y, Sikaris K, Streichert T, Macri J. Intervals on behalf of IC on R, limits (C-RIDL) D. Distinguishing reference intervals and clinical decision limits – a review by the IFCC Committee on Reference Intervals and Decision Limits. *Crit Rev Clin Lab Sci* 2018;55:420–31.
35. Anderson GJ, Frazer DM. Current understanding of iron homeostasis. *Am J Clin Nutr* 2017;106:1559S–66S.
36. Raverot V, Bonjour M, Abeillon du Payrat J, Perrin P, Roucher-Boulez F, Lasolle H, et al. Age- and sex-specific TSH upper-limit reference intervals in the general French population: there is a need to adjust our actual practices. *J Clin Med* 2020;9:792.
37. Schnabl K, Chan MK, Gong Y, Adeli K. Closing the gaps in paediatric reference intervals: the CALIPER initiative. *Clin Biochem Rev* 2008;29:89–96.
38. CLSI and IFCC. EP28-A3C; defining, establishing, and verifying reference intervals in the clinical laboratory, 3rd ed.; 2010:72 p.
39. Hoq M, Matthews S, Karlaftis V, Burgess J, Cowley J, Donath S, et al. Reference values for 30 common biochemistry analytes across 5 different analyzers in neonates and children 30 days to 18 years of age. *Clin Chem* 2019;65:1317–26.
40. Pavord S, Myers B, Robinson S, Allard S, Strong J, Oppenheimer C, et al. UK guidelines on the management of iron deficiency in pregnancy. *Br J Haematol* 2012;156:588–600.
41. Flores CJ, Yong A, Knights E, Grech J, Saxon B, Markus C, et al. Maternity iron, anaemia and blood management in South Australia: a practice-based evidence for clinical practice improvement. *Vox Sang* 2020;115:735–44.

42. Randall DA, Patterson JA, Gallimore F, Morris JM, McGee TM, Ford JB. The association between haemoglobin levels in the first 20 weeks of pregnancy and pregnancy outcomes. *PLoS One* 2019; 14:e0225123.
43. Gonzales GF, Steenland K, Tapia V. Maternal hemoglobin level and fetal outcome at low and high altitudes. *Am J Physiol-Regul Integr Comp Physiol* 2009;297:R1477–85.
44. Muñoz M, Peña-Rosas JP, Robinson S, Milman N, Holzgreve W, Breyman C, et al. Patient blood management in obstetrics: management of anaemia and haematinic deficiencies in pregnancy and in the post-partum period: NATA consensus statement. *Transfus Med* 2018;28:22–39.
45. Crispin P, Stephens B, McArthur E, Sethna F. First trimester ferritin screening for pre-delivery anaemia as a patient blood management strategy. *Transfus Apher Sci* 2019;58:50–7.
46. Daru J, Allotey J, Peña-Rosas JP, Khan KS. Serum ferritin thresholds for the diagnosis of iron deficiency in pregnancy: a systematic review. *Transfus Med* 2017;27:167–74.
47. MMWR. Recommendations to prevent and control iron deficiency in the United States. [Internet]; 1998. Available from: <https://www.cdc.gov/mmwr/preview/mmwrhtml/00051880.htm> [Accessed 08 Oct 2020].
48. World Health Organization (WHO). Serum ferritin concentrations for the assessment of iron status and iron deficiency in populations. [Internet]. Available from: https://www.who.int/vmnis/indicators/serum_ferritin.pdf [Accessed 08 Oct 2020].
49. Broek VD, Letsky W, Shenkin. Iron status in pregnant women: which measurements are valid? *Br J Haematol* 1998;103: 817–24.