Serum ferritin as an indicator of iron status: what do we need to know?

Jahnavi Daru,[1] Katherine Colman,[2] Simon J Stanworth,[3] Barbara De La Salle,[4] Erica M Wood,[2] and Sant-Rayn Pasricha[5]

[1]Women's Health Research Unit, Barts and the London School of Medicine, Queen Mary University of London, London, United Kingdom; [2]Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia; [3]NHS Blood and Transplant/Oxford University Hospitals NHS Trust, Oxford, United Kingdom; [4]United Kingdom National External Quality Assessment Service, General Haematology, Watford, United Kingdom; and [5]Medical Research Council Human Immunology Unit, Medical Research Council Weatherall Institute of Molecular Medicine, University of Oxford, United Kingdom

ABSTRACT

Determination of iron status in pregnancy and in young children is essential for both clinical and public health practice. Clinical diagnosis of iron deficiency (ID) through sampling of bone marrow to identify the absence of body iron stores is impractical in most cases. Serum ferritin (SF) concentrations are the most commonly deployed indicator for determining ID, and low SF concentrations reflect a state of iron depletion. However, there is considerable variation in SF cutoffs recommended by different expert groups to diagnose ID. Moreover, the cutoffs used in different clinical laboratories are heterogeneous. There are few studies of diagnostic test accuracy to establish the sensitivity and specificity of SF compared with key gold standards (such as absent bone marrow iron stores, increased intestinal iron absorption, and hemoglobin response to SF) among noninflamed, outpatient populations. The limited data available suggest the commonly recommended SF cutoff of <15 μg/L is a specific but not sensitive cutoff, although evidence is limited. Data from women during pregnancy or from young children are especially uncommon. Most data are from studies conducted >30 y ago, do not reflect ethnic or geographic diversity, and were performed in an era for which laboratory methods no longer reflect present practice. Future studies to define the appropriate SF cutoffs are urgently needed and would also provide an opportunity to compare this indicator with other established and emerging iron indexes. In addition, future work would benefit from a focus on elucidating cutoffs and indexes relevant to iron adequacy. *Am J Clin Nutr* 2017;106 (Suppl):1634S–9S.

Keywords: iron status, iron deficiency, ferritin, iron deficiency anemia, cutoffs, hepcidin, diagnostic test

BACKGROUND

Indicators of iron status span an array of measures and can be confounded by factors ranging from inflammation to analytic challenges. Moreover, given that iron status is a continuum from iron deficiency anemia (IDA) [i.e., reduced hemoglobin in red blood cells (RBCs)] to iron deficiency (ID) (i.e., depleted iron stores) to iron overload, different indexes may be more useful than others depending on the interest. Available indicators for these conditions include concentrations of hemoglobin, serum ferritin (SF), soluble transferrin receptor (sTfR), zinc protoporphyrin, reticulocyte hemoglobin, serum iron, and hepcidin as well as total iron-binding capacity or transferrin saturation (TSAT). Others have addressed the role and nature of these indicators (1–3). It is notable that the correlation of iron indexes with longer-term functional outcomes, such as suboptimal child development or birth weight (4), from observational cohorts or from baseline indexes from interventional trials is limited, and the lack of data to make these linkages has been the topic of a recent set of reviews conducted by the United States Preventive Services Task Force (5, 6).

Although many indexes are available, determination of status by using SF concentrations is the most commonly deployed strategy used in clinical and public health settings (7). Ferritin is an iron storage protein, regulated post-transcriptionally by cellular iron status via iron-responsive elements in its messenger RNA. Thus, higher intracellular iron concentrations result in increased ferritin expression, whereas ID inhibits expression (8). However, ferritin is also an acute-phase protein, and serum concentrations are increased in conditions of inflammation (9). During liver damage, ferritin leaks from hepatocytes, and plasma concentrations rise. The ferritin measurable in the serum appears to be chiefly derived from macrophages (10) and does not contain storage iron but reflects overall storage iron and ferritin concentrations in the liver and other tissues (11).

Thus, SF concentration is a routinely available indicator with well-described associations with iron status but also recognized limitations associated with distortions in the setting of concomitant inflammation and liver disease (12, 13). Because SF concentrations are measured along a continuous scale, the definition of SF cutoffs to determine status necessitates trade-offs between its utility as a screening tool and as a confirmatory

Presented at the workshop "Iron Screening and Supplementation in Iron-Replete Pregnant Women and Young Children" held by the NIH Office of Dietary Supplements, Bethesda, MD, 28–29 September 2016.

Supported by the Cooley's Anemia Foundation (to S-RP).

JD and KC contributed equally to this work.

Address correspondence to S-RP (e-mail: pasricha.s@wehi.edu.au).

Abbreviations used: AUC^{ROC}, area under the receiver operating characteristic curve; ID, iron deficiency; IDA, iron deficiency anemia; RBC, red blood cell; SF, serum ferritin; sTfR, soluble transferrin receptor; TSAT, transferrin saturation.

First published online October 25, 2017; doi: https://doi.org/10.3945/ajcn.117.155960.

test. The limited data available suggest the commonly recommended SF cutoff of <15 μg/L is a specific but not sensitive threshold (14–18), although evidence is limited. As pointed out by Pfeiffer and Looker (3) elsewhere in these proceedings, SF is a sensitive indicator until body iron stores are depleted, but concentrations <12 μg/L are not indicative of the severity of ID. On the other hand, sTfR concentrations are a sensitive indicator after body iron stores are depleted, and concentrations increase with increasing ID (3). However, sTfR concentrations are also a critical indicator of erythropoietic activity, limiting its use for detection of iron status if erythropoiesis is suppressed (19) or enhanced (20). Several studies have indicated that the use of the logarithm of the ratio of sTfR to SF may be a promising indicator of iron status, and usage of this ratio has become more frequent (3).

This review, conducted in support of the workshop discussions, addresses principles for developing an appropriate case definition for iron status—notably deficiency—based on a variety of indicators, evaluates evidence from relevant studies comparing SF cutoffs to key gold standards, and proposes a strategy for improving these definitions. Several organizations have been evaluating indicators to define iron status over recent years, including the WHO, which has undertaken a multifaceted approach to consider the role of SF to define ID and overload in individuals and populations (11, 21). However, there has not yet been a coordinated effort to improve the quality of underlying primary evidence.

STATISTICAL APPROACHES TO DEFINING A CUTOFF

It is important to define some of the key epidemiologic and statistical approaches that can be used to research the performance of a diagnostic test and that must be considered to appraise the quality of the existing studies. Studies evaluating diagnostic tests should usually be cross-sectional, diagnostic-test-accuracy studies (22). Participants require measurement for disease with the use of both the gold standard (which is considered a perfect dichotomous discriminator of presence or absence of disease) along with the index test (the diagnostic performance of which is under study). The sensitivity of an index test is the proportion of individuals with the disease who would be detected by the test at a particular cutoff when a continuous indicator is used, i.e., the true positives (23). When a continuous biomarker is used to dichotomously define the presence or absence of a condition, there are of course an infinite number of cutoffs that could theoretically be used. The performance of the test can be examined by plotting the true-positive (sensitivity) against false-negative (1-specificity) rate at each cutoff—a receiver operating characteristic curve (24). The area under this receiver operating characteristic curve (AUC^{ROC}) can be used to evaluate the overall performance of the test—a higher AUC^{ROC} (for example, approaching 1.0) indicates a better test, whereas a lower AUC^{ROC} (for example, approaching 0.5) indicates a poorer test (**Figure 1**) (25).

Unless a test perfectly reflects the gold standard, selection of any cutoff is a tradeoff between sensitivity and specificity. The extent of this compromise may reflect the overall AUC^{ROC} as well as biological properties of the test and condition. A cutoff with higher sensitivity may be more appropriate for screening a population. Cutoffs that simultaneously optimize sensitivity and specificity (for example, the Youden J statistic) may be useful but can also compromise both parameters (26). An additional consideration is the positive and negative likelihood ratio, which

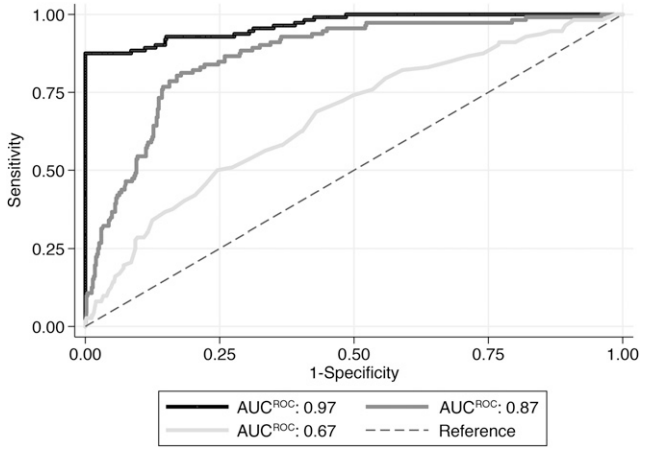

FIGURE 1 Analysis of diagnostic test accuracy. Example of a receiver operating characteristic curve; a higher AUC^{ROC} indicates a superior test for identifying the corresponding gold standard. AUC^{ROC}, area under the receiver operating characteristic curve.

provides information on the change in probability a patient has of having the disease once the test is applied (27). Thus, defining a threshold for a test is not entirely empiric. Consideration of epidemiologic and clinical factors is needed, as well as how the test will be used in the overall process of diagnosing a patient or population (28).

An alternative approach to defining abnormal thresholds is the evaluation of population reference ranges. For example, understanding the distribution of values of an indicator in a population and accepting the central 90% or 95% with a normal distribution as "healthy" may help identify individuals for whom, when a result outside this range is returned, could be considered abnormal and for whom therefore further clinical consideration is needed. Although laboratories may not formally indicate that a result outside these ranges represents disease, it is likely that many clinicians may interpret the result that way. A key limitation to this approach is that it does not specifically link the indicator to clinical evidence of disease. With this approach, 5% of the patients will be defined as abnormal regardless of biologic significance of the result.

GOLD STANDARDS FOR DETECTION OF IRON STATUS

Gold standard definitions of ID remain complex. There are several ways in which ID could be definitively considered to exist. Classically, ID is defined when examination of bone marrow aspirate under microscopy with the use of an iron stain (Perl's stain) reveals an absence of hemosiderin. This standard reflects an absence of iron available to the bone marrow for erythropoiesis, thereby resulting in anemia.

Among anemic individuals, an alternative gold standard definition is a hemoglobin response to iron treatment. This recognizes that IDA is amenable to and responds to iron treatment. This approach is often used to diagnose IDA in children, but it requires a follow-up blood test, assumes a high adherence to iron treatment, and does not account for impaired absorption of oral iron in conditions in which intestinal function is impaired by luminal disease or systemic inflammation.

Iron absorption is elevated in individuals with ID. Measurement of erythrocyte iron incorporation by using stable isotopes

(for example, Fe^{57} and Fe^{58}) is an accurate and safe measurement of iron utilization, which is enhanced in ID because of suppression of the iron regulatory hormone hepcidin. The proportion of dietary iron incorporated into the RBC mass can be estimated by measurement of the incorporation into the RBCs 2 wk after the administration of an accurate dose of a stable isotope. This technique has been used extensively to measure effects of different dietary, comorbid, and physiologic conditions on iron absorption.

Detection of iron stores with the use of one of the gold standards described above is impractical in the routine, single clinic visit and is costly and invasive. Thus, peripherally measurable indexes, such as SF concentration, are needed to identify ID.

EVIDENCE FOR CURRENT SF CUTOFFS

Diagnosis of ID with the use of SF concentration is dependent on comparing patient results with established diagnostic cutoffs. However, SF cutoffs currently recommended by several expert organizations demonstrate considerable inconsistency (**Table 1**). This is exemplified by our survey of 208 laboratories participating in the United Kingdom National External Quality Assessment Service (K Coleman, E Wood, B De La Salle, S Stanworth, and S-R Pasricha, unpublished results, 2016), which showed marked variation in SF cutoffs used to define ID in men, women, and children (**Figure 2**).

Studies are needed to clarify the utility of SF concentrations to detect ID in apparently otherwise healthy individuals or populations (rather than unwell, hospitalized patients). Although many studies have evaluated the diagnostic properties of SF in comparison with bone marrow iron stores, most of these studies were conducted in populations at high risk of inflammation, for example, in hospital populations (36–38). Such studies would likely distort the identified SF cutoff (likely raising it) because of the high prevalence of inflammation in these persons. Relatively few studies have been conducted in apparently well community populations. These are discussed below.

Hallberg et al. (14) performed a cross-sectional study that included bone marrow samples from 203 women aged 38 y in Goteberg, Sweden. The participants were part of a larger cohort study of 1462 women, and sample collection was conducted in the year 1968. Serum samples were collected simultaneously with bone marrow samples and stored at −20°C. Twenty-four years later, SF analysis was performed by using a radioimmunoassay. By comparing SF concentrations obtained from the test to those obtained when SF was assayed with the use of an early assay 14 y prior, the authors concluded there had been a 19% deterioration in SF activity in the samples over ≥14 y (and did not indicate the overall deterioration in 24 y). Using these data, the authors identified an optimal sensitivity and specificity (of 75% and 98%, respectively) at an SF concentration <16 μg/L; an SF concentration <30 μg/L resulted in a sensitivity and specificity of 93% and 75%, respectively (14).

Harju et al. (39) studied bone marrow iron stores in outpatients diagnosed with gastritis and peptic ulcer disease, a condition that is unlikely to cause systemic inflammation unless complications are present, although *Helicobacter pylori* was not measured in these patients. This study compared SF concentrations between patients with bone marrow iron stores considered "deficient," "sufficient," or "plenty." At SF concentrations <15 and <30 μg/L, sensitivity was 72% and 92%, respectively, and specificity was 96% and 92%, respectively (39). Sorbie et al. (17) measured bone marrow iron and SF concentrations in 20 healthy students as controls for a study in patients with renal failure. The study found that an SF concentration <40 μg/L had a sensitivity of 100% and specificity of 92%. Examination of the plots demonstrated that an SF concentration <15 μg/L had a sensitivity of 57% and specificity of 100% (17). Milman et al. (16) measured bone marrow iron deficiency and SF concentrations in 53 healthy students, and found that an SF concentration <15 μg/L had a sensitivity of 60% and specificity of 100%, whereas a cutoff of <30 μg/L had a sensitivity of 100% and specificity of 89%.

Specific to studies in pregnancy, the most widely used indicators of iron stores to assess iron status during pregnancy are SF and hemoglobin concentrations. There are challenges with interpretation of laboratory results related to the physiologic effects of pregnancy, and specific studies in pregnancy are required. Considering these physiologic variations, investigators have attempted to identify cutoffs defining ID by observing how bone

TABLE 1
Examples of recommended SF cutoffs to determine ID[1]

Organization	Population group	SF cutoff, μg/L
WHO (29, 30)	Children <5 y of age	<12, if inflammation <30
	Adults	<15
CDC (31)	Persons >6 mo of age	<15
Royal College of Pathologists of Australasia (32)	Adults	<30
	Prepubescent children	<20
British Society for Standardization in Haematology (2)	Pregnancy	<30
Group for the Research and Education on Anemia Therapy in Women (33)	Women	<30
Royal Australasian College of Physicians (34)	Children 1–5 y of age	<10
	Breastfed children <1 y	<5–9
	Formula-fed children <1 y	<14–39
American Association of Blood Banks (35)	Female blood donors with hemoglobin 12–12.5 g/dL	<26

[1] ID, iron deficiency; SF, serum ferritin.

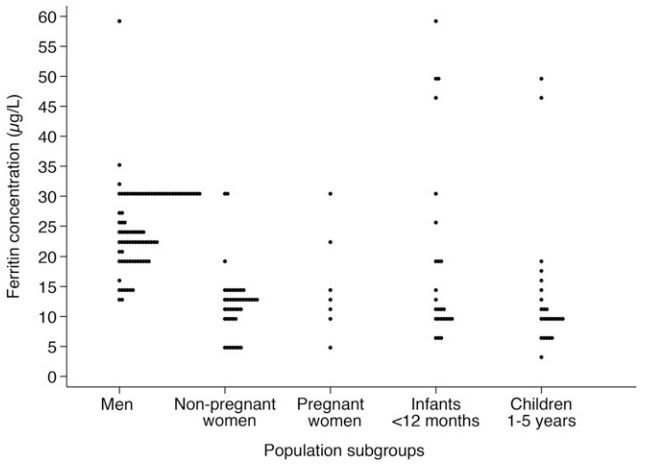

FIGURE 2 Distribution of the lower limit of the SF range in laboratories served by the UK NEQAS. In 2016 we contacted 606 laboratories predominantly in the United Kingdom and the rest of Europe participating in the UK NEQAS program and received data from 208. Laboratories were asked to provide information on the lower limit of the SF concentration as included with their results. Data on SF cutoffs used in each population category were collected. Each dot represents a laboratory reporting a specific SF cutoff in the defined population subgroup (K Coleman, E Wood, B De La Salle, S Stanworth, and S-R Pasricha, unpublished results, 2016). SF, serum ferritin; UK NEQAS, United Kingdom National External Quality Assessment Service.

marrow iron stores respond to the demands of pregnancy (14, 18). The study by van den Broek et al. (18) is the only one that has evaluated SF and bone marrow iron in pregnancy. Performed in Malawi in 1993, 47% of the study population was positive for HIV, and the mean C-reactive protein concentration was 40 mg/L, although the participants were apparently healthy. Moreover, the level of control of HIV infection was not determined (N van den Broek, Liverpool School of Tropical Medicine, personal communication, 2016). The majority of women included in the study by van den Broek et al. (18) were in the third trimester of pregnancy, when physiologic changes in inflammation and plasma volume expansion plateau. The reported sensitivity and specificity of an SF concentration <15 mg/L were 37.5% and 93.7%, respectively, and the sensitivity and specificity for an SF concentration <30 mg/L were 90% and 85%, respectively (18). Data for HIV-negative women were not separately reported.

Regarding studies in children, the study by Jonker et al. (15) is the only one evaluating SF cutoffs against bone marrow iron samples in children. This well-designed study collected bone marrow aspirates from 87 apparently healthy Malawian children aged 6–66 mo while they were anesthetized and undergoing elective orthopedic surgery. Children were excluded if they had clinical evidence of inflammation. In this study, the sensitivity and specificity of an SF concentration <12 μg/L were 44.7% and 89.6%, respectively. For an SF concentration <18 μg/L these values were 73.7% and 77.1%, respectively, and for an SF concentration <30 μg/L these values were 81.6% and 37.5%, respectively. This was the only study to include a receiver operating characteristic analysis, and it reported an AUC^{ROC} for SF to detect absent bone marrow stores of 0.797 (essentially identical the SF/sTfR index for which AUC^{ROC} was 0.801, *P*-difference = 0.90) (15).

Together, these studies demonstrate that evidence to support any recommended SF cutoff for diagnosis of ID is limited. The commonly reported threshold of 15 μg/L is likely specific but can be expected to miss many cases of ID—perhaps as many as half. An SF concentration cutoff <30 μg/L is associated with a higher sensitivity but more false-positive diagnoses. Importantly, there are very limited data to support SF thresholds in the critical groups of pregnant women and young children. The data are predominantly derived from older studies, with only one study published in the last 20 y (15). Furthermore, the laboratory methodology used in these studies is likely considerably different from current practice, especially for reference standards and commutable calibration materials. Finally, there is a paucity of data from ethnically and geographically diverse settings.

One consideration for interpretation of SF cutoffs is the pretest probability of an individual patient or population for having a disease. Where the pretest probability is high, a more sensitive cutoff may be appropriate because a larger proportion of negative results will be false negatives; conversely, when the probability of disease is low, most positive results will be false positives, and hence perhaps a lower cutoff, which is more specific, could be considered. For example, a higher SF cutoff might be appropriate when seeking to confirm a diagnosis of ID in an anemic individual compared with in an individual who is undergoing routine screening. Likewise, in biological states in which iron is rapidly used, such as among pregnant women or growing children (40), a higher (not lower) SF cutoff might be more appropriate.

ALTERNATIVE APPROACHES TO DEFINING SF CUTOFFS

SF reference ranges

SF reference ranges have been proposed by measuring the distribution of values in a population and defining abnormality as outside the central 95% of the population. However, this approach does not directly measure a biological relation between SF concentration and iron stores. Moreover this approach is influenced by the prevalence of ID and inflammation (or other factors that influence SF) in the study population. Thus, SF concentrations will be lower in premenopausal women and children compared with men because of menstruation and growth, respectively. The effect of these differences is demonstrated by diagnosis of ID at a lower cutoff in women than in men; if the biological relation between SF (adjusted for inflammation) and bone marrow iron stores (and functional outcomes) is constant regardless of sex and age, then SF cutoffs should not differ between these groups. Defining ID by using SF reference ranges must, therefore, make careful consideration of the underlying population.

Iron absorption

Several studies demonstrate a close inverse correlation between baseline SF concentration and erythrocyte iron incorporation. ID is associated with an increase in erythrocyte iron incorporation that can be measured by using stable isotope studies (41). Hicks et al. (42) demonstrated correlations of $r = 0.64$ and 0.60 between SF and Fe^{57} utilization in 5- to 6- and 9- to 10-mo-old children, respectively. If a definition of increased iron absorption were applied (for example, >20%), these

relations could be used to define ID. This approach has been used to define cutoffs of hepcidin concentration to detect ID (43).

Hemoglobin response to iron supplementation

Increases in hemoglobin concentrations after anemic individuals are treated with iron may indicate that the anemia is secondary to ID (for example, defined as a 1- to 2-g/dL increment in hemoglobin concentrations over 2–4 wk) (44). A study among anemic nonpregnant Vietnamese women found that weekly iron supplementation for 12 wk resolved anemia in 56% of cases. A baseline SF concentration <15 μg/L had a 44% sensitivity and 80% specificity for predicting a response to iron (defined as a 1-g/dL increase in hemoglobin or cure of anemia, seen in 66% of women), whereas an SF concentration <30 μg/L had a sensitivity of 72% and specificity of 52% (19). The optimal simultaneous sensitivity and specificity for predicting a "cure" of anemia were seen at an SF concentration of 26 μg/L.

CONSIDERATIONS RELATIVE TO DETERMINING IRON ADEQUACY AND CONDITIONS OF IRON REPLETION

Although iron indexes have historically been used chiefly to establish the presence of ID clinically and at the public health level, there is increasing interest in also identifying individuals and populations who can be considered iron replete. As discussed above, existing cutoffs for SF (for example, <15 μg/L) are specific but poorly sensitive for ID, whereas higher cutoffs (for example, <30 μg/L) are more sensitive, but the literature is limited. Hallberg et al. (14) demonstrated that 75% of iron-replete individuals had SF concentrations >30 μg/L, with only 7% of iron-deficient individuals having an SF concentration above this cutoff. However, in children, Jonker et al. (15) showed that only 37.5% of Malawian children with bone marrow iron repletion had an SF concentration >30 μg/L, indicating this cutoff may be too high to diagnose iron repletion among this population. Further studies remain essential to characterize the diagnostic properties of SF to detect iron repletion. Older indicators of iron status, such as TSAT and serum iron, are distorted by inflammation but may be useful in uninflamed individuals to identify a normal supply of iron to the bone marrow and tissues; elevated TSAT remains perhaps the best screening test of iron overload (45). Because it is distorted by inflammation, liver disease, and obesity, SF is not a useful screening test for iron overload but may be more useful to stage the severity of iron loading once the diagnosis has been established. Alternative approaches to detecting iron repletion may include the measurement of hepcidin, which could indicate hepatic sensing of adequate iron stores and hence homeostatic limitation of iron absorption.

CONCLUSIONS AND FUTURE DIRECTIONS

New diagnostic-test-accuracy studies are needed to inform cutoffs. Such studies would ideally involve a cross-sectional design in a community setting with the use of a population free from inflammation, with sampling from multiple centers, countries, and regions and of different ages, sexes, and ethnicities. Particular attention to pregnant women and children is needed. These studies should include bone marrow iron and erythrocyte iron incorporation measured by using stable isotopes. Although retrieval of bone marrow aspirates from community-dwelling healthy individuals seems challenging, creative study designs can achieve this. For example, appropriate samples are already routinely collected during harvesting of bone marrow for allogeneic-stem cell transplantation, and collection of simultaneous peripheral blood samples would complement this. Likewise, bone marrow aspirates could be collected from noninflamed individuals undergoing elective surgical procedures, especially orthopedic procedures. Furthermore, stable isotope studies are safe and achievable in both children and in pregnancy (46). These studies could simultaneously compare a broad range of iron indexes, enabling an optimal diagnostic approach to be defined. In addition to diagnostic test accuracy studies comparing SF concentrations to physiological variables of ID, correlation of iron indexes with longer-term functional outcomes could help define the implications of a diagnosis of ID. However, such data alone cannot define cutoffs for ID but rather would help confirm that the ID is associated with critical health outcomes.

There is variation in recommendations for SF cutoffs indicative of ID among different expert organizations and even laboratories, which impairs the development of recommendations to implement screening programs. This complicates clinical guidelines for diagnosis and treatment of ID, and obscures meaningful epidemiologic assessment of the burden of this condition. There is an urgent need to undertake further primary research to develop an evidence base for cutoffs of iron indexes defining ID and other states of iron status.

The authors' responsibilities were as follows—JD, KC, and S-RP: drafted the manuscript; KC, SJS, BDLS, EMW, and S-RP: undertook the survey of laboratories; and all authors: read and approved the final manuscript. S-RP has been employed as a consultant to the WHO. None of the other authors reported a conflict of interest related to this study.

REFERENCES

1. Pasricha SR, Flecknoe-Brown SC, Allen KJ, Gibson PR, McMahon LP, Olynyk JK, Roger SD, Savoia HF, Tampi R, Thomson AR, et al. Diagnosis and management of iron deficiency anaemia: a clinical update. Med J Aust 2010;193:525–32.
2. Pavord S, Myers B, Robinson S, Allard S, Strong J, Oppenheimer C; British Committee for Standards in Haematology. UK guidelines on the management of iron deficiency in pregnancy. Br J Haematol 2012;156:588–600.
3. Pfeiffer CM, Looker AC. Laboratory methodologies for indicators of iron status: strengths, limitations, and analytical challenges. Am J Clin Nutr 2017;106(Suppl):1606S–14S.
4. Pratt JJ, Khan KS. Non-anaemic iron deficiency - a disease looking for recognition of diagnosis: a systematic review. Eur J Haematol 2016;96:618–28.
5. McDonagh M, Blazina I, Dana T, Cantor A, Bougatsos C. Routine iron supplementation and screening for iron deficiency anemia in children ages 6 to 24 months: a systematic review to update the U.S. Preventive Services Task Force recommendation. Rockville (MD): Agency for Healthcare Research and Quality; 2015.
6. McDonagh M, Cantor A, Bougatsos C, Dana T, Blazina I. Routine iron supplementation and screening for iron deficiency anemia in pregnant women: a systematic review to update the U.S. Preventive Services Task Force recommendation. Rockville (MD): Agency for Healthcare Research and Quality; 2015.
7. Mei Z, Cogswell ME, Parvanta I, Lynch S, Beard JL, Stoltzfus RJ, Grummer-Strawn LM. Hemoglobin and ferritin are currently the most efficient indicators of population response to iron interventions: an analysis of nine randomized controlled trials. J Nutr 2005;135:1974–80.
8. Muckenthaler MU, Galy B, Hentze MW. Systemic iron homeostasis and the iron-responsive element/iron-regulatory protein (IRE/IRP) regulatory network. Annu Rev Nutr 2008;28:197–213.

9. Elin RJ, Wolff SM, Finch CA. Effect of induced fever on serum iron and ferritin concentrations in man. Blood 1977;49:147–53.
10. Cohen LA, Gutierrez L, Weiss A, Leichtmann-Bardoogo Y, Zhang DL, Crooks DR, Sougrat R, Morgenstern A, Galy B, Hentze MW, et al. Serum ferritin is derived primarily from macrophages through a non-classical secretory pathway. Blood 2010;116:1574–84.
11. Garcia-Casal MN, Pasricha S-R, Martinez RX, Lopez-Perez L, Peña-Rosas JP. Serum or plasma ferritin concentration as an index of iron deficiency and overload (Protocol). Cochrane Database Syst Rev 2015; 7:CD011817.
12. Blake DR, Waterworth RF, Bacon PA. Assessment of iron stores in inflammation by assay of serum ferritin concentrations. Br Med J (Clin Res Ed) 1981;283:1147–8.
13. Dale JC, Burritt MF, Zinsmeister AR. Diurnal variation of serum iron, iron-binding capacity, transferrin saturation, and ferritin levels. Am J Clin Pathol 2002;117:802–8.
14. Hallberg L, Bengtsson C, Lapidus L, Lindstedt G, Lundberg PA, Hulten L. Screening for iron deficiency: an analysis based on bone-marrow examinations and serum ferritin determinations in a population sample of women. Br J Haematol 1993;85:787–98.
15. Jonker FA, Boele van Hensbroek M, Leenstra T, Vet RJ, Brabin BJ, Maseko N, Gushu MB, Emana M, Kraaijenhagen R, Tjalsma H, et al. Conventional and novel peripheral blood iron markers compared against bone marrow in Malawian children. J Clin Pathol 2014;67:717–23.
16. Milman N, Bangsboll S, Strandberg Pedersen N, Visfeldt J. Serum ferritin in non-dialysis patients with chronic renal failure: relation to bone marrow iron stores. Scand J Haematol 1983;30:337–44.
17. Sorbie J, Valberg LS, Corbett WE, Ludwig J. Serum ferritin, cobalt excretion and body iron status. Can Med Assoc J 1975;112:1173–8.
18. van den Broek NR, Letsky EA, White SA, Shenkin A. Iron status in pregnant women: which measurements are valid? Br J Haematol 1998; 103:817–24.
19. Pasricha SR, Casey GJ, Phuc TQ, Mihrshahi S, MacGregor L, Montresor A, Tien N, Biggs BA. Baseline iron indices as predictors of hemoglobin improvement in anemic Vietnamese women receiving weekly iron-folic acid supplementation and deworming. Am J Trop Med Hyg 2009;81:1114–9.
20. Pasricha SR, Hayes E, Kalumba K, Biggs BA. Effect of daily iron supplementation on health in children aged 4-23 months: a systematic review and meta-analysis of randomised controlled trials. Lancet Glob Health 2013;1:e77–86.
21. Garcia-Casal MN, Peña-Rosas JP, Pasricha S-R. Rethinking ferritin cutoffs for iron deficiency and overload. Lancet Haematol 2014;1:e92–4.
22. Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane handbook for systematic reviews of diagnostic test accuracy version 100: the Cochrane collaboration. London: The Cochrane Collaboration; 2009.
23. Kirkwood BR, Sterne JAC. Essential medical statistics. Oxford (United Kingdom): Blackwell Science; 2003.
24. Mallett S, Halligan S, Thompson M, Collins GS, Altman DG. Interpreting diagnostic accuracy studies for patient care. BMJ 2012;345:e3999.
25. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
26. Fluss R, Faraggi D, Reiser B. Estimation of the Youden index and its associated cutoff point. Biom J 2005;47:458–72.
27. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ 2004; 329:168–9.
28. Man-Son-Hing M, Laupacis A, O'Connor AM, Coyle D, Berquist R, McAlister F. Patient preference-based treatment thresholds and recommendations: a comparison of decision-analytic modeling with the probability-tradeoff technique. Med Decis Making 2000;20:394–403.
29. WHO. Serum ferritin concentrations for the assessment of iron status and iron deficiency in populations. Geneva (Switzerland): Vitamin and Mineral Nutrition Information System; 2011.
30. WHO, CDC. Assessing the iron status of populations: including literature reviews. Report of a Joint World Health Organization/Centers for Disease Control and Prevention Technical Consultation on the assessment of iron status at the population level, Geneva, Switzerland, 6–8 April 2004. 2nd ed. Geneva (Switzerland): WHO/CDC; 2007.
31. Recommendations to prevent and control iron deficiency in the United States. MMWR Recomm Rep 1998;47:1–29.
32. RCPA. Iron studies standardised reporting protocol. Surry Hills (New South Wales): Royal College of Pathologists of Australasia; 2013.
33. Friedman AJ, Shander A, Martin SR, Calabrese RK, Ashton ME, Lew I, Seid MH, Goodnough LT. Iron deficiency anemia in women: a practical guide to detection, diagnosis, and treatment. Obstet Gynecol Surv 2015;70:342–53.
34. Grant CC, Wall CR, Brewster D, Nicholson R, Whitehall J, Super L, Pitcher L. Policy statement on iron deficiency in pre-school-aged children. J Paediatr Child Health 2007;43:513–21.
35. AABB. Management of risk for iron deficiency in female blood donors with HB levels of 12.0-<12.5 g/dL using ferritin testing [Internet]. [cited 2016 Oct 4]. Available from: http://www.aabb.org/advocacy/regulatorygovernment/donoreligibility/hsim/Pages/Management-of-Risk-for-Iron-Deficiency-in%20Female-Blood-Donors-with-HB-Levels-12.0-12.5.aspx.
36. Guagnozzi D, Severi C, Ialongo P, Viscido A, Patrizi F, Testino G, Vannella L, Labriola R, Strom R, Caprilli R. Ferritin as a simple indicator of iron deficiency in anemic IBD patients. Inflamm Bowel Dis 2006;12:150–1.
37. Ong KH, Tan HL, Lai HC, Kuperan P. Accuracy of various iron parameters in the prediction of iron deficiency in an acute care hospital. Ann Acad Med Singapore 2005;34:437–40.
38. Witte DL, Angstadt DS, Davis SH, Schrantz RD. Predicting bone marrow iron stores in anemic patients in a community hospital using ferritin and erythrocyte sedimentation rate. Am J Clin Pathol 1988;90: 85–7.
39. Harju E, Pakarinen A, Larmi T. A comparison between serum ferritin concentration and the amount of bone marrow stainable iron. Scand J Clin Lab Invest 1984;44:555–6.
40. Scholl TO, Reilly T. Anemia, iron and pregnancy outcome. J Nutr 2000;130:443S–7S.
41. Prentice AM, Doherty CP, Abrams SA, Cox SE, Atkinson SH, Verhoef H, Armitage AE, Drakesmith H. Hepcidin is the major predictor of erythrocyte iron incorporation in anemic African children. Blood 2012;119:1922–8.
42. Hicks PD, Zavaleta N, Chen Z, Abrams SA, Lönnerdal B. Iron deficiency, but not anemia, upregulates iron absorption in breast-fed peruvian infants. J Nutr 2006;136:2435–8.
43. Pasricha SR, Atkinson SH, Armitage AE, Khandwala S, Veenemans J, Cox SE, Eddowes LA, Hayes T, Doherty CP, Demir AY, et al. Expression of the iron hormone hepcidin distinguishes different types of anemia in African children. Sci Transl Med 2014;6:235re3.
44. Schrier SL, Auerbach M. Treatment of iron deficiency anemia in adults. In: Post TW, ed. UpToDate. Waltham (MA): Wolters Kluwer; 2016.
45. Bacon BR, Adams PC, Kowdley KV, Powell LW, Tavill AS. American Association for the Study of Liver D. Diagnosis and management of hemochromatosis: 2011 practice guideline by the American Association for the Study of Liver Diseases. Hepatology 2011;54:328–43.
46. Whittaker PG, Lind T, Williams JG. Iron absorption during normal human pregnancy: a study using stable isotopes. Br J Nutr 1991;65: 457–63.