Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature

Mengke Du

Date

Iron Biomarker Profiles and Prediction of Iron Deficiency Anemia among Women of Reproductive Age in the US

By

Mengke Du

Master of Public Health

Department of Biostatistics and Bioinformatics

Yi-An Ko, PhD
(Thesis Advisor)

Hanqi Luo, PhD
(Reader)

Melissa Fox Young, PhD
(Reader)

Iron Biomarker Profiles and Prediction of Iron Deficiency Anemia among Women of Reproductive Age in the US

By

Mengke Du

Master of Public Health

Bachelor of Science

Central South University

2020

Thesis Committee Chair: Yi-An Ko, PhD

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Biostatistics

2023

Abstract

Iron Biomarker Profiles and Prediction of Iron Deficiency Anemia among Women of Reproductive Age in the US

By Mengke Du

Iron deficiency (ID) and iron deficiency anemia (IDA) are significant public health concerns among women of reproductive age in the United States (Looker 1997a). Serum ferritin (SF) is considered the gold standard for assessing ID (Daru et al. 2017) and, when combined with hemoglobin (Hb), is used to diagnose IDA. However, SF is not routinely measured in clinical practice. This study aimed to evaluate the prevalence of ID using nine iron biomarkers, including SF, and to assess the performance of alternative biomarkers for predicting IDA in non-pregnant women aged 12-49 years using data from the National Health and Nutrition Examination Survey (NHANES) (n = 12781). Serum iron had the highest estimated prevalence of ID (31.58%), while soluble transferrin receptor (sTfR) had the lowest (2.21%), and SF showed a prevalence of 26.42%. Stratified by age, race, parity, and family poverty income ratio, the estimates revealed a higher prevalence of ID among non-Hispanic Black populations, those with lower family parity income ratios, and younger age groups (12-19 years old). Cluster analysis identified three subgroups of non-pregnant women with different patterns of iron biomarkers and prevalence of IDA, with Cluster 3 having the highest prevalence (61.95%). We developed a model to predict IDA using RBC and MCV, achieving an area under the receiver operating characteristic (ROC) Curve (AUC) of 98.61% in the training set and a sensitivity of 97.4%, and a specificity of 93.7% in the test set. These findings highlight the heterogeneity of ID and the importance of tailored approaches to diagnose IDA. Further research is needed to validate the prediction model and develop effective interventions to prevent and diagnose IDA in non-pregnant women.

Iron Biomarker Profiles and Prediction of Iron Deficiency Anemia among Women of Reproductive Age in the US

By

Mengke Du

Master of Public Health

Bachelor of Science

Central South University

2020

Thesis Committee Chair: Yi-An Ko, PhD

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Biostatistics

2023

Acknowledgments

I am deeply grateful to my thesis advisors, Yi-An Ko and Hanqi Luo, for their invaluable guidance and support throughout my thesis project. Yi-An's technical insights enabled me to dive into my research questions and overcome the challenges of data analysis. Hanqi's expertise in the field of nutrition has been crucial in helping me explore and formulate my thesis question effectively. Meanwhile, they offered constructive feedback that enhanced my writing and communication skills. In addition to their academic guidance, their warm and approachable demeanor made them feel like supportive mentors rather than distant professors, fostering my personal growth and deepening my grasp of the discipline.

I would also like to extend my sincere gratitude to the Rollins School of Public Health at Emory University for giving me the opportunity to pursue my graduate studies. The supportive and collaborative learning environment at Emory has been invaluable, and I have been privileged to access top-notch facilities and resources that facilitated my research. I am particularly grateful for the guidance and mentorship of faculty and staff in the BIOS department, who consistently encouraged and supported me throughout this journey.

Lastly, I want to say a huge thank you to my family and friends for always having my back throughout my academic journey. Their love, encouragement, and belief in me have kept me going, even during tough times. I'm so thankful for everything they've done for me. As I continue my path, their support inspires me to keep pushing forward with determination and passion.

Table of contents

1 Introduction

ID is a significant public health concern that affects individuals of all ages and genders worldwide. According to the World Health Organization (WHO), over 2 billion people globally suffer from anemia, and ID is the most common cause of anemia (World Health Organization 2020). In the United States, approximately 10% of women of reproductive age have ID, and up to 5% have IDA (Looker 1997a). The diagnosis of ID and IDA is essential to prevent adverse health outcomes, including impaired cognitive function, decreased work productivity, and poor pregnancy outcomes (World Health Organization 2020).

Serum ferritin (SF) is the most commonly used biomarker for assessing ID and low SF concentrations reflect a state of iron depletion (Daru et al. 2017). However, SF levels can be affected by inflammation and liver disease, leading to false-positive results (World Health Organization 2011). Soluble transferrin receptor (sTfR) is another biomarker for ID that reflects the concentration of the transferrin receptor in the bloodstream. sTfR levels are elevated in ID, making it a useful marker for detecting early ID. The sTfR/log ferritin ratio, known as total body iron (TBI), is also a reliable biomarker for diagnosing ID, particularly in the early stages (Skikne et al. 2011).

Serum iron, total iron-binding capacity (TIBC), and transferrin saturation (TS) are valuable in distinguishing between nutritional ID and anemia caused by chronic infections, inflammation, or neoplastic diseases (Gibson 2005). TIBC measures the total amount of iron that transferrin can bind, and serum iron measures the concentration of iron in the bloodstream. However, TIBC and serum iron have limited sensitivity and specificity for detecting ID (World Health Organization 2011). Furthermore, TS can be calculated from serum iron and TIBC levels and reflects the percentage of iron-binding sites on transferrin that is saturated with iron. TS is a useful biomarker for assessing iron status, but it has limitations and should be interpreted along with other iron biomarkers such as SF and sTfR (Beard and Tobin 2000).

Various red blood cell parameters, such as mean corpuscular volume (MCV), red blood cell distribution width (RDW), and red blood cell count (RBC), are also used as iron biomarkers. MCV and RDW are indirect measures of erythropoiesis and are used to assess the quality of red blood cells. RBC measures the number of red blood cells in circulation, which can be decreased in ID. These biomarkers have unique mechanisms for reflecting iron status and can help identify different stages of ID.

As a common type of anemia, IDA occurs when ID is sufficiently severe to reduce erythropoiesis. The diagnosis of IDA is based on a low hemoglobin (Hb) level in the presence of ID, and currently, SF and hemoglobin are commonly used for IDA screening. Combining SF with other biomarkers, such as sTfR, MCV, RDW, and RBC, has been shown to improve the accuracy of diagnosing IDA (Goddard et al. 2011). However, the lack of routine availability of SF testing and the inconsistent use of iron biomarkers in research highlight the need for comparison of these iron biomarkers in detecting ID, as well as alternative ways to diagnose IDA,

This study aims to provide valuable insights into ID prevalence based on different iron biomarkers and efficient IDA diagnosis. Firstly, we compared different iron biomarkers, including SF, sTfR, TS, TIBC, serum iron, TBI, MCV, RDW, and RBC, in estimating the prevalence of ID and IDA among women of reproductive age in the United States. Secondly, the study explored the development of alternative methods to diagnose IDA when SF testing is not readily available, which could lead to more effective screening and management of these conditions. The findings from this research have the potential to significantly impact the clinical approach to diagnosing and managing ID and IDA globally.

2 Subjects and methods

2.1 Study population and sample selection

National Health and Nutrition Examination Survey (NHANES) is a multi-purpose survey aimed at evaluating the health and nutritional status of both adults and children in the United States (National Center for Health Statistics, CDC 2023a). Currently, it has been conducted continuously since 1999 and

involves a household interview followed by a standardized physical examination at a Mobile Examination Center. The survey uses a stratified probability sampling method to select participants from different counties, blocks, households, and individuals within those households. It is carried out by the National Center for Health Statistics (NCHS) and the Centers for Disease Control and Prevention (CDC). Approval from the NCHS Research Ethics Review Board was obtained and written informed consent was gathered from participants over the age of 12 and parental consent was obtained for those under 18. The methods for data collection and laboratory methods have been documented on NHANES website (National Center for Health Statistics, CDC 2023b).

We combined NHANES 1999 - 2010 and 2015 - 2018 data for women aged 15 - 49 years. We excluded NHANES 2011-2014 since C-reactive protein (CRP) is not available. The reproductive status of subjects was based on urine pregnancy test, self-reported pregnancy, and dual-energy x-ray absorptiometry (DXA) examination. We restricted our study sample to non-pregnant women only (n = 12781). For the purpose of the present first part of analysis, sample weights for NHANES 1999-2000 were based on population estimates developed by the Census Bureau before the Year 2000 Decennial Census counts became available. The two-year sample weights for NHANES 2001-2002, and all other subsequent two-year cycles, are based on population estimates that incorporate the year 2000 Census counts (National Center for Health Statistics, CDC 2023c). For variables used in the analysis, we use the 4-year weights provided by NCHS for 1999-2002, then include the 2-year weights for each additional 2-year cycle that is combined. In the second part of analysis, we did not take into account the original sample weights assigned to the data because we did not intend to perform population-based analyses. we also excluded those who had missing SF, sTfR, Hb, RBC, MCV, Serum iron, TS, or TIBC (n = 8213).

2.2 Laboratory method and variable definition

CRP was quantified by latex-enhanced nephelometry method and assays were performed on a Behring Nephelometer for quantitative CRP determination in 1999-2010. In 2015-2016, CRP was measured on the Beckman Coulter UniCel DxC 600 Synchron and the Beckman Coulter UniCel DxC 660i Synchron

Access chemistry analyzers Previous analyses indicated that no statistical adjustment is needed for results obtained between the two Beckman UniCel® analyzers. During the 2017-2018 cycle the Roche Cobas 6000 chemistry analyzer (Cobas 6000) was used for the entire 2017-2018 cycle. For highly sensitive CRP levels above 20 mg/L based on Cobas 6000 (or 23 mg/L based on DxC 660i), the relationship between the two instruments appeared opposite that of the lower values, with 2017-2018 measurements higher than 2015-2016 measurements (citation). A weighted Deming regression was used to adjust the CRP (mg/L) result:

- $Cobas\,6000 = 0.8695 \times DxC\,660i + 0.2954$

SF was measured using several methods over years. In 1999-2002, Ferritin is measured by using the Bio-Rad Laboratories' "QuantImune Ferritin IRMA" kit (Bio-Rad Laboratories, 1986), which is a single-incubation two-site immunoradiometric assay (IRMA) based on the general principles of assays. (Addison et al., 1972). A same method was used in 2003 but this assay was discontinued by the manufacturer in early 2004, so ferritin was measured by the Roche Tina-quant® Ferritin immunoturbidimetric assay on the Hitachi 912 clinical analyzer (Roche Diagnostics, Indianapolis, IN, USA) in 2004-2008. The Hitachi method gave higher ferritin results than the Bio-Rad method. Due to the difference, three piecewise linear regression equations were used to adjust the 2003 SF (ng/mL) data to be comparable to the 2004 data:

- $SF \leq 25: Hitachi = 1.2534 \times Bio - Rad + 1.4683$
- $25 < SF \leq 65: Hitachi = 1.2001 \times Bio - Rad + 1.4693$
- $SF > 65: Hitachi = 1.0791 \times Bio - Rad + 4.8183$

This was accomplished prior to the data release by NCHS. To adjust the 1999-2002 data to compare with the 2003-2004 data, the same regression equations were used. The Roche Elecsys-170 sandwich immunoassay was used in 2009-2010. To convert the Roche E170 SF (ng/mL) concentrations to be equivalent to the Hitachi method, we applied a conversion equation following NHANES analytic notes:

- $Hitachi = 10^{\frac{(log_{10}(E170) - 0.049)}{0.989}}$

The Roche Mod E170 analyzer method was used for most of 2015-2016 SF data and replaced with the Roche Cobas e601 analyzer in mid-2016 and later in 2017-2018. On average, ferritin values measured from the Roche e601 analyzer were 8.8% higher than values from the Roche Mod E170 ($p<.0001$). A weighted Deming regression was used to adjust the e601 SF results (ng/mL) to be comparable with the E170 result:

- $E170 = -0.2079 + e601 \times 0.9271$

Then the same conversion equation in 2009-2010 was used to keep consistency. After calibration SF values, we applied Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) regression approach to adjust SF for inflammation. The reproducibility of the BRINDA method was published in detail elsewhere.

The method for measurement of sTfR is immuno-turbidimetry using Roche kits on the Hitachi 912 clinical analyzer in 1999-2008 and on the Hitachi Mod P clinical analyzer 2009-2010. In 2015-2016, the method principle for measurement of sTfR is a particle enhanced immunoturbidimetric assay that uses Roche kits on the Roche Mod P analyzer and Cobas® c501 clinical analyzer. The Roche Mod P analyzer was used for most of 2015-2016 and replaced with the Roche Cobas c501 analyzer in mid-2016 and later in 2017-2018. No adjustment of the data was needed.

TBI (mg/kg) was calculated as previously described in detail (citation) from sTfR and SF concentrations by using a formula:

- $TBI\ stores = -[log_{10}(sTfR \times 1000 \div SF) - 2.8229] \div 0.1207$

Serum iron and TIBC was measured by a modification of the automated AAII-25 colorimetric method in 1999-2000, which is based on the procedures of Giovaniello et al. and of Ramsey (citation). In 2001-

2010 and 2015-2016 the method used to measure the iron concentration was a timed-endpoint method and the serum iron was measured on the Beckman/Coulter LX20 analyzer and Beckman UniCel® DxC800 analyzer, respectively. TIBC was calculated indirectly using the unsaturated iron binding capacity (UIBC) method in 2001-2006 and not available in 2007-2010 and 2015-2016. In 2017-2018, serum iron was measured using a three-step process with FerroZine reagent on Roche Cobas 6000. TIBC was calculated indirectly using Iron (frozen), serum and UIBC. TS was calculated as $(iron/TIBC) \times 100\%$ in 1999-2006 and 2017-2018. TIBC was not available in 2007-2010 and 2015-2016.

The methods used to derive complete blood count (CBC) parameters (Hb, MCV, RBC, RDW) are based on the Beckman Coulter® method of counting and sizing, in combination with an automatic diluting and mixing device for sample processing, and a single beam photometer for hemoglobinometry.

Demographic characteristics were based on self-reported data. Age was stratified into three groups: 15-19, 20-34 and 35-49 years. Race-ethnicity following the NHANES guidelines included Mexican American, Other Hispanic, Non-Hispanic White, Non-Hispanic Black, and other. Parity was based on self-report of the number of pregnancies that resulted in a live birth in women aged ≥12 y who reported that they had ever been pregnant. Parity was categorized into 0, 1, or ⩾2 births. Family poverty income ratio, which is an index that measures the ratio of family income to poverty with poverty being defined by the Department of Health and Human Services' (HHS) poverty guidelines, was categorized the family poverty income ratio as < 130% or ⩾130%. This cutoff is commonly used to determine eligibility for various government assistance programs (US Department of Agriculture 2022).

There are different cutoff points for abnormal values of anemia and ID of indicators stated above. WHO has different hemoglobin standards for different ages and sex groups. For nonpregnant women, the cutoff point is <120 g/L and is <110 g/L for pregnant women (Looker 1997b). WHO has set up a global cutoff for hemoglobin levels, but the cut-off points might not be applicable for special populations. One study examined hemoglobin distributions of healthy White, Black, Mexican, Hispanic, and Asian non-pregnant

women to check race-specific hemoglobin cut-offs for mild anemia. They found that the Asian cut-off was lower than the WHO standard at 11.22g/dL (Varghese, Thomas, and Kurpad 2019). RBC for females was considered normal in the range of 4.2 – 5.8 million cells/μL and for males, 3.6 – 5.6 million cells/μL. The cut-off values for the red blood cell indices were as follows: MCV < 81 fL for people 15-49 years old (Mei et al. 2003), and RDW < 14.5 % (Abdelrahman et al. 2012). For serum iron and TIBC, the cutoff points are 60 μg/dL (10.74 μmol/L) and 410 μg/dL (73.39 μmol/L), respectively (Cook and Finch 1979). For transferrin saturation < 16% a subject is considered to be ID (Hallberg, The Swedish Nutrition Foundation, and The Swedish Society of Medicine 1996). The cut-off value of SF is < 15 μg/L for nonpregnant women (Anon 2020). The cutoff value of sTfR is >8.3 mg/L for adults (Phiri et al. 2009). And the cutoff value of TBI is < 0 mg/kg (Cook, Flowers, and Skikne 2003).

2.3 Statistical analysis

Exam sample weights, primary sampling units (PSUs), and strata were used for analyses. We applied "survey" package in R (4.2.0) written by Thomas Lumley to derive statistics. For some PSUs that had only one unit, we specified "survey.lonely.psu = 'adjust'" to accurate estimation of population parameters.

We defined ID as having abnormal value of the following 9 tests: SF, sTfR, TBI, MCV, RBC, RDW, Serum iron, TS and TIBC. To be considered iron deficient, an individual had to have an abnormal value for any one of these biomarkers (Table 1). For descriptive analysis, we visualized ID prevalence according to each iron biomarker and further stratified by age groups (12–19, 20–34, and 35–49 y), race-ethnic group (Mexican American, Other Hispanic, Non-Hispanic White, Non-Hispanic Black and other), parity (0, 1, and >=2) and family income poverty income ratio (<130% compared with ⩾130%), Next, we calculated pairwise Pearson's correlations between iron biomarkers. Furthermore, for those who had Hb, SF, sTfR, TBI, MCV, RBC, RDW, Serum iron, TS, and TIBC available, we repeated the visualization procedure among individuals with anemia.

Table 1. Abnormal values of different iron biomarkers

Iron biomarker	Acronym	Abnormal values
Serum ferritin*	SF	< 15 µg/L
Serum iron	serum iron	< 10.74 µmol/L
Transferrin saturation	TS	< 16%
Total body iron stores†	TBI	< 0 mg/kg
Mean corpuscular volume	MCV	< 81 fL
Red blood cell count	RBC	< 4.2 million cells/µL
Red blood cell distribution width	RDW	> 14.5 %
Total iron binding capacity	TIBC	> 4100 µg/L
Soluble transferrin receptor	sTfR	> 8.3 mg/L

* SF is adjusted using the BRINDA regression correction method (Namaste et al. 2017)
† $TBI\ stores\ = -[log10\ (sTfR\ \times\ 1000\ \div\ adj\_SF)\ -\ 2.8229]\ \div\ 0.1207$

Since we did not intend to perform population-based studies, the sample weights were not included in this part of analysis. IDA was defined as having abnormal values for Hb and SF, and was treated IDA as an outcome. First, in order to identify the optimal number of clusters for our analysis, we used the Elbow method, Silhouette method, and Gap statistics. K-means clustering was used for partitioning the participants into different clusters with distinct iron biomarker profiles, considering 8 iron biomarkers (sTfR, Hb, RBC, MCV, Serum iron, TS, and TIBC). Hierarchical clustering was used to verify the results of K-means. Subsequently, logistic regression was used to examine the association between the clusters and the probability of having IDA and if the association is modified by demographic characteristics by including interactions.

All the above iron biomarkers, except TBI (since TBI is a function of SF and sTfR), were considered as candidate predictors for IDA. The model performance was measured by Area Under the Curve (AUC) and the accessibility of iron predictors. First, we split our data into a training set (n=2741, 60%) and a test set (n=1827, 40%). The training set was used to fit the logistic regression model and selected the model with the most effective combinations of iron biomarkers. we fitted logistic regression models with all combination sets of predictors (127 in total). For the selected model, we assessed the relationship between IDA and iron predictors adjusting for demographic covariates (age, race, parity, and family poverty income ratio). We further assessed the interaction between the iron predictors in the selected model. We used ROC curves to characterize the sensitivity and specificity of screening for IDA. And we derived a threshold using the point on the ROC curves which maximized the Youden's index. In the test set, we calculated a fitted possibility of having IDA for each individual using the selected model and classified them as "IDA" or "not IDA" using the derived threshold. Then we presented sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for all individuals. We also repeated this procedure in the test set for people who were identified having anemia by Hb < 120 g/L for further validation.

3 RESULTS

In the total sample of US non-pregnant women from 1999 to 2010 and 2015-2018, the proportions of US non-pregnant women who were aged 15-19, 20-34, and 35-49 y were 10.45%, 43.69%, and 45.86%, respectively. The proportions of US non-pregnant women who were Mexican American, Other Hispanic, Non-Hispanic White, Non-Hispanic Black and other were 9.80%, 6.86%, 62.54%, 13.35%, and 7.45% respectively. The proportions of US pregnant women who were in 0, 1, and ≥2 parity at the time of the medical examination were 34.40%, 17.34%, and 48.26%, respectively. For 26.29% of US non-pregnant women, the family income was below the 130% poverty income ratio (Table 2).

Table 2. Sample size and weighted proportion by group

	Group	Sample size	Proportion (%)
Age	15 - 19	2972	10.45
	20 - 34	4947	43.69
	35 - 49	4862	45.86
Race	Mexican American	2953	9.80
	Other Hispanic	1015	6.86
	Non-Hispanic White	4676	62.54
	Non-Hispanic Black	3048	13.35
	Other Race	1089	7.45
Parity	0	3868	34.40
	1	1644	17.34
	>= 2	4526	48.26
Poverty income ratio	< 130%	4239	26.29
	>= 130%	7482	73.71

Prevalence of ID in US non-pregnant women estimated different iron biomarkers was showed in Figure 1. The highest estimated prevalence of ID was obtained using serum iron among the 9 iron biomarkers analyzed (31.58%). The other biomarkers that showed an estimated prevalence of ID in decreasing order were TS, SF, TIBC, RBC, RDW, MCV, TBI. The lowest estimated prevalence of ID was obtained using sTfR (2.21%).

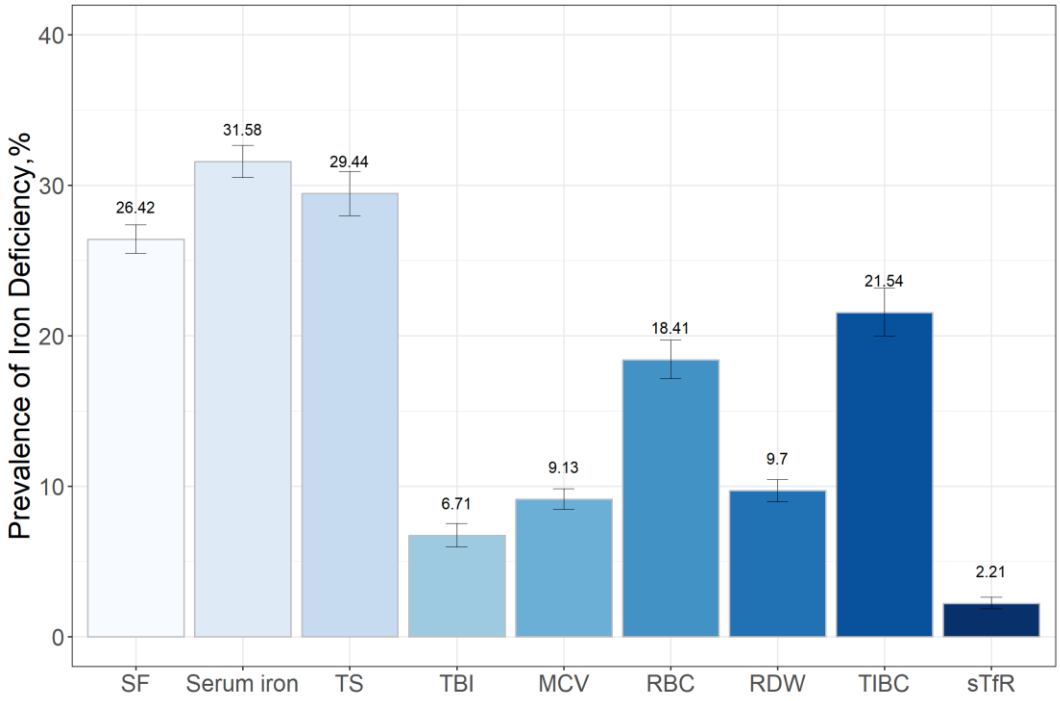

Figure 1. Prevalence of iron deficiency estimated by different iron biomarkers

When estimating the prevalence of ID using different iron biomarkers, it's not surprising that stratifying these estimates by additional factors, such as age(15-19, 20-34, and 35-49 y), race (Mexican American, Other Hispanic, Non-Hispanic White, Non-Hispanic Black and other), parity (0, 1, and ≥ 2), and poverty income ratio(<130% and ≥ 130%), resulted in further differences in the estimated prevalence of IDA. For example, in general, younger age group (15-19 y) had a higher prevalence of ID compared to other age groups. We found that there was a higher prevalence of ID among non-Hispanic Black populations compared to non-Hispanic White populations. Non-pregnant women who had more parity were at a higher risk of having ID. The lower family parity income ratio was associated with an increased prevalence of ID. It's worth noting that even within these subgroups, there were similarities in the higher and lower pattern of estimated IDA prevalence between different iron biomarkers. The highest estimated prevalence of ID was obtained using serum iron among the 9 iron biomarkers in most of subgroups and the lowest estimated prevalence of ID was obtained using sTfR, still (Figure 2).

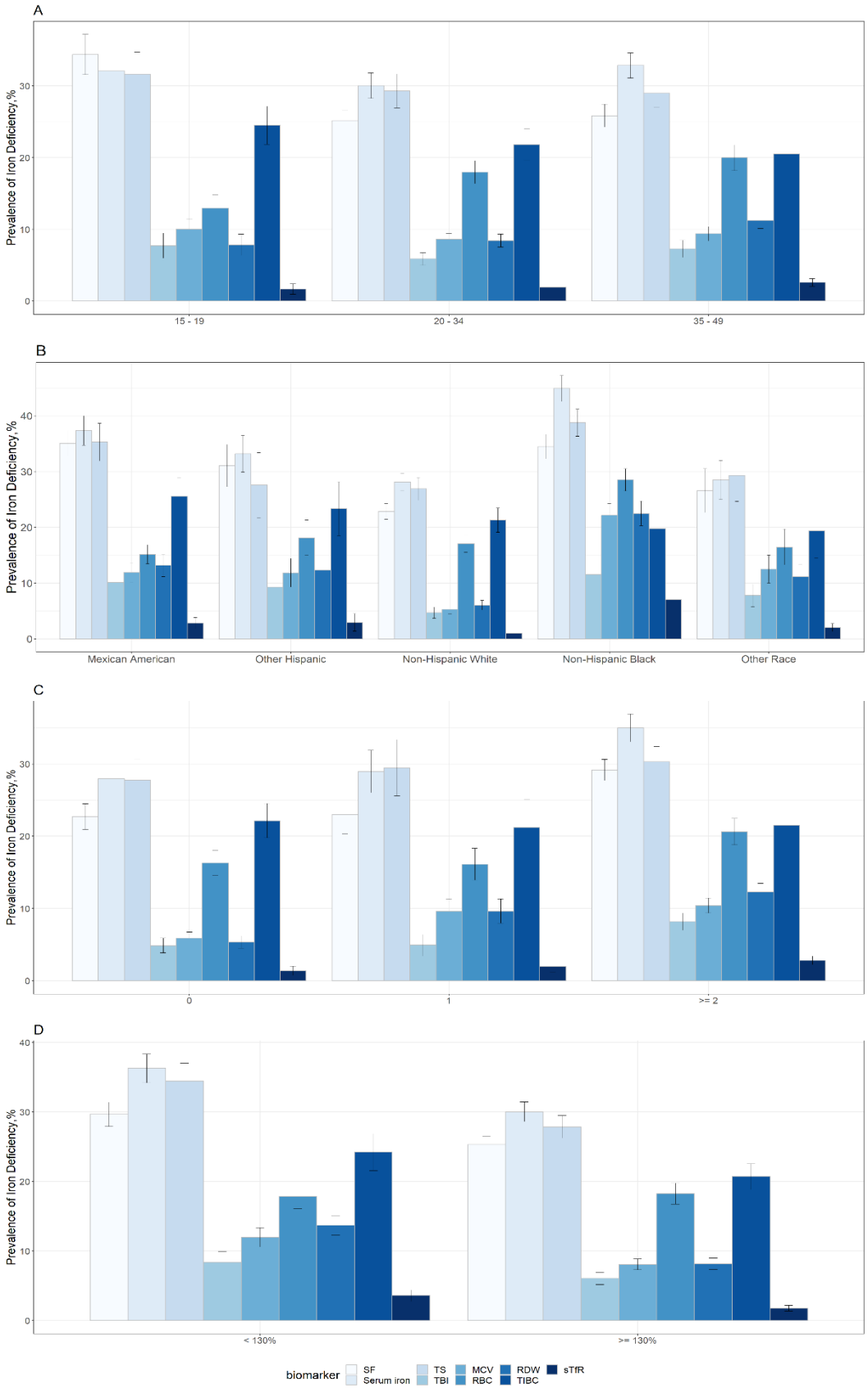

Figure 2. Prevalence of iron deficiency estimated by different iron biomarkers among different age (A),

race(B), parity(C), and poverty income ratio (D) groups

Overall, there was no strong correlation between any two of these biomarkers except Serum iron and TS. It was not surprising since TS was calculated using serum iron and TIBC (Figure 3).

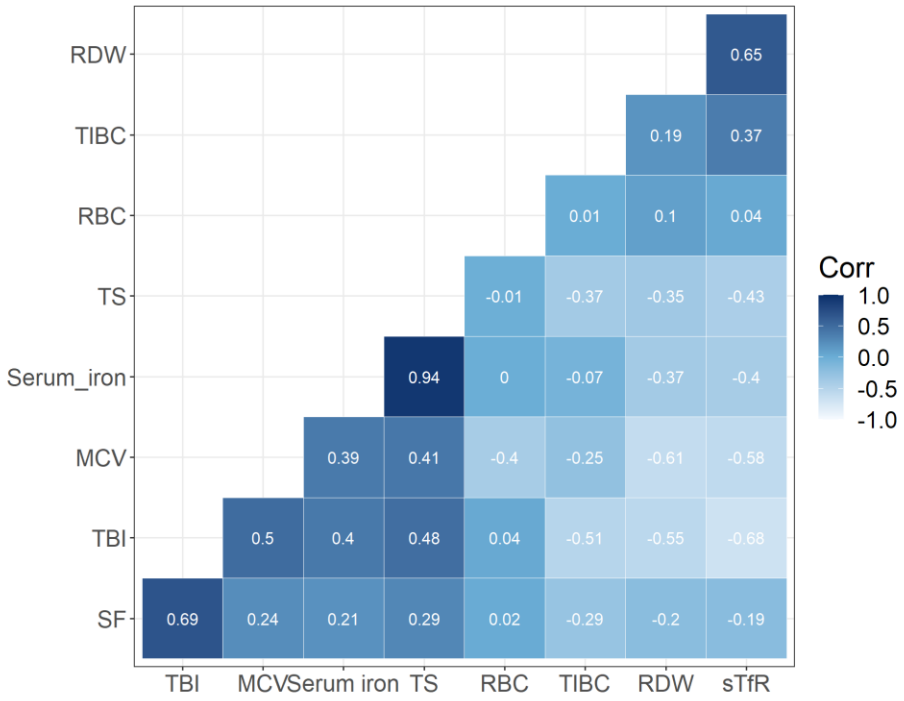

Figure 3. Pairwise correlation between two biomarkers among all iron biomarkers

Figure 4 shows the proportion of women with ID estimated by different biomarkers among those with anemia (IDA). When considering individuals with both hemoglobin and all iron biomarkers available, the prevalence of ID estimated by SF, serum iron, and TS was high and similar, ranging from 74.26% to 77.04%. On the other hand, the prevalence of ID estimated by TBI, MCV, RBC, RDW, TIBC, and sTfR was relatively lower, ranging from 21.94% to 57.53%.

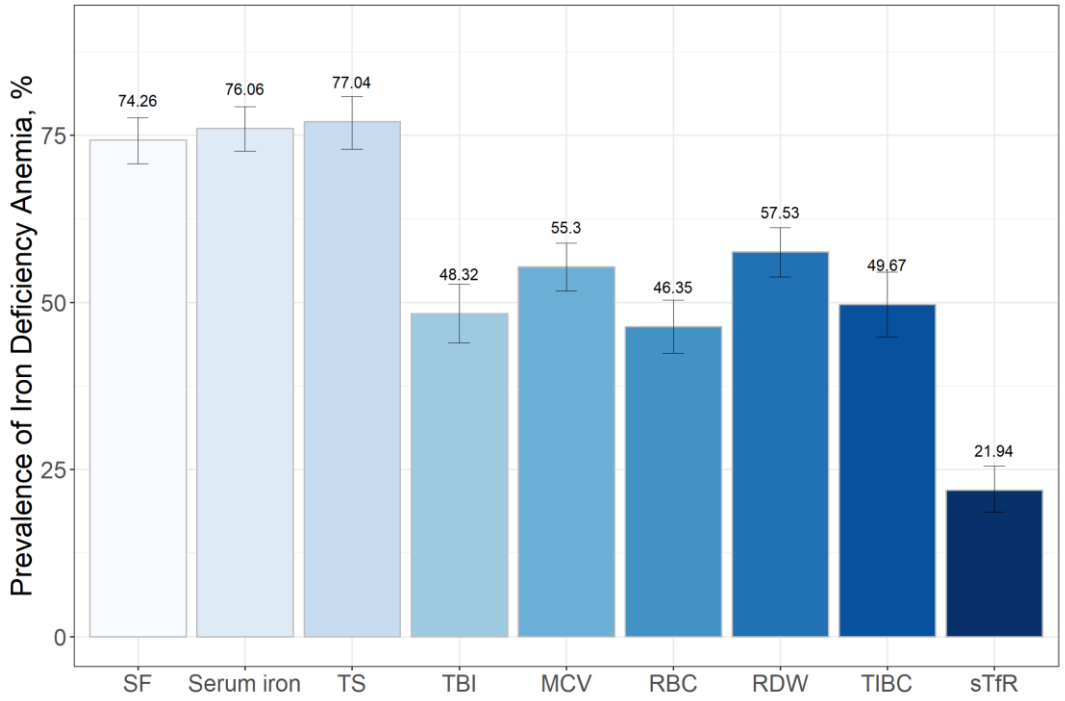

Figure 4. Prevalence of iron deficiency estimated by different iron biomarkers among anemic women

After excluding those who had missing SF, sTfR, Hb, RBC, MCV, Serum iron, TS, or TIBC, final sample included size of 4568. Three clusters represented distinct subgroups with different patterns of iron biomarkers and prevalence of IDA. We found that the standardized mean value of RBC was similar among all three clusters. Cluster 1 had higher standardized mean values for Serum iron, TS, TBI, and MCV and relatively lower standardized mean values for RDW, TIBC, and sTfR. On the other hand, Cluster 3 had lower standardized mean values for serum iron, TS, TBI, and MCV, and higher standardized mean values for RDW, TIBC, and sTfR. The values of iron biomarkers for Cluster 2 were average, indicating a normal population. Additionally, we found that Cluster 3 had the highest prevalence (61.95%) of IDA, followed by Cluster 2 (2.99%) and Cluster 1 (0.51%) (Figure 5).

Figure 5. standardized mean of different iron biomarkers across three cluster groups (left); Prevalence of IDA of each cluster (right)

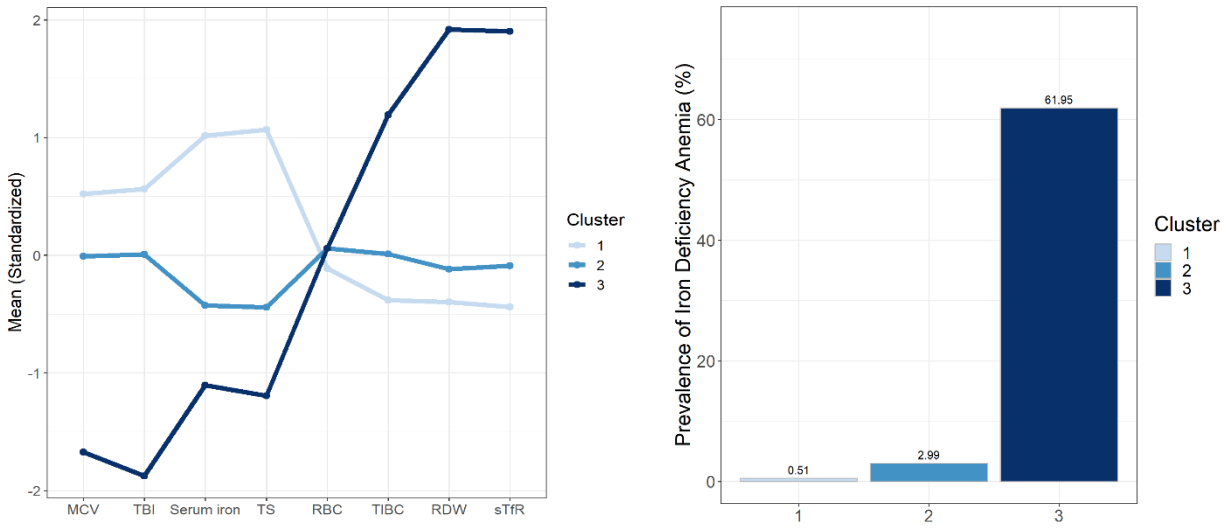

Figure 5. standardized mean of different iron biomarkers across three cluster groups (left); Prevalence of IDA of each cluster (right)

Logistic regression suggested the relationship between IDA risk and clusters remains significant after controlling for covariates (age, race, parity, and family poverty income ratio) (Table 3). And there were no significant interactions between clusters and demographic characteristics.

Characteristic	OR[1]	95% CI[1]	p-value
cluster3			
2	—	—	
1	248.17	122.20, 596.12	**<0.001**
3	4.64	2.25, 11.22	**<0.001**
Age(year)			
15 - 19	—	—	
20 - 34	0.74	0.43, 1.26	0.3
35 - 49	0.86	0.47, 1.56	0.6
Race/Ethnicity			
Non-Hispanic White	—	—	
Mexican American	0.89	0.53, 1.50	0.7
Other Hispanic	1.90	0.82, 4.18	0.12
Non-Hispanic Black	3.27	2.13, 5.08	**<0.001**
Other Race	2.01	1.05, 3.79	**0.033**
Parity			
0	—	—	
1	1.46	0.83, 2.55	0.2
>= 2	1.48	0.87, 2.54	0.2
Family income: poverty			
>= 130%	—	—	
< 130%	1.09	0.76, 1.55	0.6

[1] OR = Odds Ratio, CI = Confidence Interval

Table 3. Parameters estimates for logistic regression model

We identified an optimal model for predicting IDA that included RBC and MCV, with an AUC of 98.61% in the training set. To determine the optimal threshold for categorizing individuals as having IDA or not, we used Youden's index, which was calculated as the maximum value of (sensitivity + specificity - 1) across all possible thresholds. The resulting threshold of 0.107 was the value that maximized the Youden's

index in our study. When applied to the test dataset, the model had a sensitivity of 97.4%, specificity of 93.7%, PPV of 58.3%, and NPV of 99.7% using the derived threshold to determine IDA status. For anemic women in the test dataset, the model had a sensitivity of 97.4%, specificity of 15.4%, PPV of 82.0%, and NPV of 60.0% (Table 4).

Table 4. Performance results of the optimal model

		IDA status			
		All women		Anemia women	
		Positive	Negative	Positive	Negative
Results from model	Positive	148	106	148	33
	Negative	4	1572	4	6
Calculation	Sensitivity	97.4%		97.4%	
	Specificity	93.7%		15.4%	
	PPV	58.3%		82%	
	NPV	99.7%		60%	

4 DISCUSSION

ID is a major public health concern worldwide, particularly among non-pregnant women. The WHO considers SF as the primary biomarker for detecting ID in the general population (World Health Organization 2007). However, our study suggests that the prevalence of ID in non-pregnant women in the United States may vary depending on the biomarker used to estimate it. We found that serum iron had the highest estimated prevalence of ID (31.58%), while sTfR had the lowest estimated prevalence (2.21%). These findings suggest that relying on a single biomarker like SFmay not accurately capture the prevalence of ID in all populations, stressing the need for further research to refine iron status assessment and biomarker selection.

Previous research has also noted the challenges and limitations associated with different iron biomarkers, including SF, when assessing iron status in populations with inflammation or infection (Ganz 2013). The NHANES dataset has been extensively used to assess iron status using SF and total body iron models, providing valuable insights into the strengths and limitations of these approaches ((Pfeiffer and Looker 2017)). Stratifying the estimates by age, race, parity, and poverty income ratio resulted in further differences in the estimated prevalence of ID, with younger age groups, non-Hispanic Black populations, and those with lower family parity income ratios having a higher prevalence of ID. These findings are consistent with previous studies and highlight the importance of considering these factors in developing targeted interventions to prevent and manage ID in at-risk populations (Barton et al. 2020; Looker 1997a).

Furthermore, our study identified distinct subgroups of non-pregnant women based on patterns of iron biomarkers and prevalence of IDA. Cluster analysis revealed three distinct subgroups with different patterns of iron biomarkers and prevalence of IDA, highlighting the heterogeneity of ID in non-pregnant women. These subgroups may help clinicians identify individuals who are at a higher risk of developing IDA and may guide the development of targeted interventions to prevent and manage this condition.

In addition to identifying distinct subgroups, we developed a model to predict IDA using RBC and MCV. The model had high sensitivity and specificity when applied to the test dataset, indicating its potential clinical usefulness. Furthermore, RBC and MCV values can be obtained from a standard complete blood count (CBC) test that is routinely ordered for many clinical indications, including anemia. This suggests that the prediction model we developed may be more feasible and cost-effective for clinicians to use in clinical practice than other biomarkers like SF. In recent years, there has been growing interest in using artificial intelligence (AI) and machine learning (ML) to predict and diagnose various health conditions, including anemia. Saputra developed an AI-based model using extreme learning machine (ELM) algorithm to predict and distinguish between different types of anemia, including beta thalassemia trait (BTT), IDA, hemoglobin E (HbE), and combination anemia (Saputra, Sunat, and Ratnaningsih 2023). The model achieved high accuracy, sensitivity, precision, and F1 score, indicating its potential clinical

usefulness in identifying and managing anemia in patients. Such predictive models could be useful in resource-limited settings, where access to laboratory testing and specialized healthcare may be limited. However, further research is needed to validate these models in different populations and settings before they can be widely implemented in clinical practice.

Several limitations of our study should be considered. Firstly, our adjustment for inflammation using CRP did not consider the potential confounding effects of acute phase proteins such as alpha-1-acid glycoprotein (AGP) and C-reactive protein (CRP) together due to the lack of CRP data. This may have led to an overestimation or underestimation of ID prevalence, particularly among subgroups with a higher prevalence of inflammation. Secondly, there is no consensus on cutoff points for iron biomarkers in categorizing ID, which can lead to varying estimates of ID prevalence and limit the generalizability of our findings. Thirdly, the sample size used to develop the model predicting IDA using RBC and MCV may not be ideal for generalization to other populations or clinical settings. Additionally, calculating an individual's IDA status based on the model may not be practical for clinicians in some resource-limited settings. Finally, our study focused on non-pregnant women of reproductive age, limiting the generalizability of our findings to other populations such as men or older adults. Future studies should explore the prevalence and patterns of ID in these populations to better understand the burden of this condition and guide the development of targeted interventions.

Our study aimed to investigate the prevalence and patterns of ID in non-pregnant women in the United States and evaluate the use of different iron biomarkers to assess iron status. By identifying distinct subgroups and developing a predictive model, we provided important insights into the accuracy of IDA diagnosis in clinical practice. However, further research is needed to address the limitations of our study and to validate the findings in other populations. Moreover, future studies should focus on developing effective interventions to prevent and manage ID, particularly in at-risk populations such as younger age groups, non-Hispanic Black populations, and those with lower family parity income ratios, who have a higher prevalence of ID. By addressing these gaps in knowledge, we can improve our understanding of

the complex mechanisms underlying ID and IDA and work towards improving the health outcomes of those affected by this condition.

REFERENCE

Abdelrahman, Esam G., Gasim I. Gasim, Imad R. Musa, Leana M. Elbashir, and Ishag Adam. 2012. "Red Blood Cell Distribution Width and Iron Deficiency Anemia among Pregnant Sudanese Women." *Diagnostic Pathology* 7:168. doi: 10.1186/1746-1596-7-168.

Anon. 2020. *WHO Guideline on Use of Ferritin Concentrations to Assess Iron Status in Individuals and Populations*. Geneva: World Health Organization.

Barton, James C., Howard H. Wiener, Ronald T. Acton, Paul C. Adams, John H. Eckfeldt, Victor R. Gordeuk, Emily L. Harris, Christine E. McLaren, Helen Harrison, Gordon D. McLaren, and David M. Reboussin. 2020. "Prevalence of Iron Deficiency in 62,685 Women of Seven Race/Ethnicity Groups: The HEIRS Study" edited by J. R. Connor. *PLOS ONE* 15(4):e0232125. doi: 10.1371/journal.pone.0232125.

Beard, John, and Brian Tobin. 2000. "Iron Status and Exercise,,." *The American Journal of Clinical Nutrition* 72(2):594S-597S. doi: 10.1093/ajcn/72.2.594S.

Cook, J. D., and C. A. Finch. 1979. "Assessing Iron Status of a Population." *The American Journal of Clinical Nutrition* 32(10):2115–19. doi: 10.1093/ajcn/32.10.2115.

Cook, James D., Carol H. Flowers, and Barry S. Skikne. 2003. "The Quantitative Assessment of Body Iron." *Blood* 101(9):3359–63. doi: 10.1182/blood-2002-10-3071.

Daru, Jahnavi, Katherine Colman, Simon J. Stanworth, Barbara De La Salle, Erica M. Wood, and Sant-Rayn Pasricha. 2017. "Serum Ferritin as an Indicator of Iron Status: What Do We Need to Know?" *The American Journal of Clinical Nutrition* 106(Suppl 6):1634S-1639S. doi: 10.3945/ajcn.117.155960.

Ganz, Tomas. 2013. "Systemic Iron Homeostasis." *Physiological Reviews* 93(4):1721–41. doi: 10.1152/physrev.00008.2013.

Gibson, Rosalind S. 2005. *Principles of Nutritional Assessment*. 2nd ed. New York: Oxford University Press.

Goddard, A. F., M. W. James, A. S. McIntyre, B. B. Scott, and on behalf of the British Society of Gastroenterology. 2011. "Guidelines for the Management of Iron Deficiency Anaemia." *Gut* 60(10):1309–16. doi: 10.1136/gut.2010.228874.

Hallberg, Leif, The Swedish Nutrition Foundation, and The Swedish Society of Medicine, eds. 1996. *Iron Nutrition in Health and Disease: The Swedish Nutrition Foundation's 20th International Symposium*. London: Libbey.

Looker, Anne C. 1997a. "Prevalence of Iron Deficiency in the United States." *JAMA: The Journal of the American Medical Association* 277(12):973. doi: 10.1001/jama.1997.03540360041028.

Looker, Anne C. 1997b. "Prevalence of Iron Deficiency in the United States." *JAMA: The Journal of the American Medical Association* 277(12):973. doi: 10.1001/jama.1997.03540360041028.

Mei, Zuguo, Ibrahim Parvanta, Mary E. Cogswell, Elaine W. Gunter, and Laurence M. Grummer-Strawn. 2003. "Erythrocyte Protoporphyrin or Hemoglobin: Which Is a Better Screening Test for Iron

Deficiency in Children and Women?" *The American Journal of Clinical Nutrition* 77(5):1229–33. doi: 10.1093/ajcn/77.5.1229.

Namaste, Sorrel ML, Fabian Rohner, Jin Huang, Nivedita L. Bhushan, Rafael Flores-Ayala, Roland Kupka, Zuguo Mei, Rahul Rawat, Anne M. Williams, Daniel J. Raiten, Christine A. Northrop-Clewes, and Parminder S. Suchdev. 2017. "Adjusting Ferritin Concentrations for Inflammation: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) Project." *The American Journal of Clinical Nutrition* 106:359S-371S. doi: 10.3945/ajcn.116.141762.

National Center for Health Statistics, CDC. 2023a. "NHANES - National Health and Nutrition Examination Survey Homepage." Retrieved April 6, 2023 (https://www.cdc.gov/nchs/nhanes/index.htm).

National Center for Health Statistics, CDC. 2023b. "NHANES Questionnaires, Datasets, and Related Documentation." Retrieved April 6, 2023 (https://wwwn.cdc.gov/nchs/nhanes/Default.aspx).

National Center for Health Statistics, CDC. 2023c. "NHANES Tutorials - Weighting Module." Retrieved April 6, 2023 (https://wwwn.cdc.gov/nchs/nhanes/tutorials/weighting.aspx).

Pfeiffer, Christine M., and Anne C. Looker. 2017. "Laboratory Methodologies for Indicators of Iron Status: Strengths, Limitations, and Analytical Challenges." *The American Journal of Clinical Nutrition* 106:1606S-1614S. doi: 10.3945/ajcn.117.155887.

Phiri, K. S., J. C. J. Calis, A. Siyasiya, I. Bates, B. Brabin, and M. B. van Hensbroek. 2009. "New Cut-off Values for Ferritin and Soluble Transferrin Receptor for the Assessment of Iron Deficiency in Children in a High Infection Pressure Area." *Journal of Clinical Pathology* 62(12):1103–6. doi: 10.1136/jcp.2009.066498.

Saputra, Dimas Chaerul Ekty, Khamron Sunat, and Tri Ratnaningsih. 2023. "A New Artificial Intelligence Approach Using Extreme Learning Machine as the Potentially Effective Model to Predict and Analyze the Diagnosis of Anemia." *Healthcare* 11(5):697. doi: 10.3390/healthcare11050697.

Skikne, Barry S., Kari Punnonen, Paul H. Caldron, Michael T. Bennett, Mari Rehu, Gail H. Gasior, Janna S. Chamberlin, Linda A. Sullivan, Kurtis R. Bray, and Paula C. Southwick. 2011. "Improved Differential Diagnosis of Anemia of Chronic Disease and Iron Deficiency Anemia: A Prospective Multicenter Evaluation of Soluble Transferrin Receptor and the STfR/Log Ferritin Index." *American Journal of Hematology* 86(11):923–27. doi: 10.1002/ajh.22108.

US Department of Agriculture. 2022. "SNAP Eligibility | Food and Nutrition Service." Retrieved April 6, 2023 (https://www.fns.usda.gov/snap/recipient/eligibility).

Varghese, JithinSam, Tinku Thomas, and AnuraV Kurpad. 2019. "Evaluation of Haemoglobin Cut-off for Mild Anaemia in Asians - Analysis of Multiple Rounds of Two National Nutrition Surveys." *Indian Journal of Medical Research* 150(4):385. doi: 10.4103/ijmr.IJMR_334_18.

World Health Organization. 2007. *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: World Health Organization.

World Health Organization. 2011. “Serum Ferritin Concentrations for the Assessment of Iron Status and Iron Deficiency in Populations. Vitamin and Mineral Nutrition Information System.”

World Health Organization. 2020. *WHO Guideline on Use of Ferritin Concentrations to Assess Iron Status in Individuals and Populations*. Geneva: World Health Organization.