Hidden Iron Deficiency in High-Performance Adolescents: A Multi-Omics and Predictive Analytics Framework

Authors

Amal AlNatour[1]
Hadeel Saed[2]

[1] Faculty of Educational Sciences, Yarmouk University, Irbid, Jordan.
[2] Department of English Language and Translation, Faculty of Arts and Humanities, Applied, Science Private University, Amman, Jordan.

Corresponding author:
Amal AlNatour
amalalnatour56@gmail.com

Abstract

Purpose: The research question of this study was to explore the multifaceted interaction between iron status, inflammation, and aerobic performance in adolescent athletes, as well as to establish a sophisticated, integrative model of early iron deficiency diagnosis and prediction of performance beyond conventional single-biomarker methods.

Methodology: The study utilized a longitudinal cohort design that included multidimensional data such as hematological (hemoglobin, ferritin), inflammatory (CRP, IL-6), physiological (VO 2max), nutritional (nutrition), and training load. Machine learning (Random Forest, XGBoost, neural networks) was used in conjunction with statistical analyses (multivariate regression, correlation analysis) to identify nonlinear relationships. To combine these variables into one predictive framework, a novel composite measure, the Iron Performance Index (IPI), was developed.

Findings: These findings revealed that hemoglobin and ferritin were major positive predictors of VO2max, while CRP and IL-6 showed significant negative associations, indicating the influence of inflammation on iron bioavailability and performance. The machine learning models had high predictive accuracy (up to 87%), high sensitivity (85.5%), and high specificity (95.2%). The most influential variables were ferritin, IL-6, VO2max, and hemoglobin, as identified by feature importance analysis. The IPI effectively measured the multidimensional reciprocity of the physiological system and more accurately identified subclinical iron deficiency.

Conclusions: The results suggest that dynamic interactions between iron metabolism and inflammation, rather than isolated biomarkers, determine aerobic performance. The Iron Performance Index and the proposed integrative framework offer a more precise, holistic, and clinically meaningful method for monitoring athletes' health and optimizing performance, representing a significant improvement over traditional diagnostic practice.

Keywords

Iron deficiency; Adolescent athletes; Aerobic performance; VO_2max; Ferritin; Hemoglobin; Inflammation; CRP; IL-6; Machine learning.

Resumen

Propósito: Analizar la interacción entre el estado del hierro, la inflamación y el rendimiento aeróbico en atletas adolescentes, y desarrollar un modelo integrador para el diagnóstico temprano de la deficiencia de hierro.

Metodología: Se empleó un diseño longitudinal con variables hematológicas, inflamatorias, fisiológicas, nutricionales y de carga de entrenamiento. Se aplicaron técnicas de aprendizaje automático (Random Forest, XGBoost, redes neuronales) junto con análisis estadísticos. Se desarrolló un índice compuesto denominado Índice de Rendimiento del Hierro (IPI).

Resultados: La hemoglobina y la ferritina mostraron asociaciones positivas con el VO_2max, mientras que la PCR y la IL-6 presentaron asociaciones negativas. Los modelos alcanzaron una alta precisión (hasta 87%), con elevada sensibilidad y especificidad. El IPI permitió identificar con mayor precisión la deficiencia subclínica de hierro.

Conclusiones: La interacción entre el metabolismo del hierro y la inflamación determina el rendimiento aeróbico. El IPI ofrece un enfoque más preciso y clínicamente útil que los métodos tradicionales.

Método más preciso, holístico y clínicamente relevante para el monitoreo de la salud de los atletas y la optimización del rendimiento, que representa una mejora significativa respecto a los enfoques diagnósticos tradicionales.

Palabras clave

Deficiencia de hierro; atletas adolescentes; rendimiento aeróbico; VO_2max; ferritina; hemoglobina; inflamación; PCR; IL-6; aprendizaje automático.

Introduction

Iron deficiency (ID) is currently one of the most widespread micronutrient disorders around the world, with more than two billion people affected by it, and adolescents constitute a particularly frail group because of rapid growth, hormonal fluctuations, and metabolic demands (Safiri et al., 2021; Leung et al., 2024). Exercise-induced physiological stressors, such as hemolysis, gastrointestinal microbleeding, mineral loss from sweating, and iron sequestration mediated by inflammation, also affect iron metabolism in athletic groups, particularly in endurance-based sports (Sim et al., 2019; Pedlar et al., 2019; Kuwabara et al., 2022). Historically, iron deficiency has been evaluated using static hematological biomarkers, including hemoglobin (Hb), serum ferritin (Fer), and serum iron (sFe) (see Figure.1). Nonetheless, recent studies indicate that these indicators might not sufficiently capture early or functional iron deficit, especially in athletes, in which hepcidin-mediated inflammatory mechanisms might conceal actual iron bioavailability (Garcia-Casal et al., 2021; Tarancon-Diez et al., 2022). This restriction is particularly acute in young athletes, as subclinical iron deficiency can impair mitochondrial function, oxygen delivery, and, eventually, aerobic efficiency without causing overt anemia.

One of the most important factors determining athletic performance in endurance sports is aerobic performance, measured by maximal oxygen uptake (VO2 max) (Tomkinson et al., 2017). Iron is a critical factor in oxygen delivery (hemoglobin), storage (myoglobin), and aerobic metabolism (cytochromes), so it cannot be ignored to achieve optimal aerobic metabolism (Heffernan et al., 2019). There is evidence that iron status is associated with VO2max, although the results are inconsistent, especially in adolescent groups, where developmental changes and training adaptations introduce greater variability (Shoemaker et al., 2019; Keller et al., 2024). The background research shows a serious paradox: even with sufficient VO2max and nutritional status, a significant percentage of adolescent athletes had low ferritin levels, indicating latent iron deficiency. Such a lack of connection highlights the need for a more nuanced interpretation of iron metabolism that goes beyond standard biomarkers and incorporates dynamic physiological, nutritional, and inflammatory variables.

The latest developments in the sports science community have suggested a systems-biology approach that combines several physiological variables, such as hematological, metabolic, and inflammatory markers, along with performance measures, to provide deeper insight into athlete health and performance (Pedlar et al., 2019; Kardasis et al., 2023). Also, machine learning methods have become powerful tools for discovering complex, non-linear correlations between biomarkers and performance outcomes, enabling the prediction of subclinical deficiencies in advance that traditional statistical models might miss. Moreover, adolescence is a physiological window marked by rapid growth, endocrine alterations, and high nutritional needs (Cohen and Powers, 2024). These factors can exacerbate iron imbalance when coupled with high training loads, especially under dietary or socio-economic restrictions (Monyeki et al., 2024). Although this is the case, most available research relies on cross-sectional designs and small panels of biomarkers, which limit its ability to elucidate dynamic relationships between iron metabolism and performance. The main goal of the research is to develop and test a data-driven model for identifying latent iron deficiency in high-performance adolescents using multi-omics data and predictive analytics. In particular, the study will (i) combine heterogeneous biological and clinical data sources into a single analytical framework, (ii) quantify the relative contribution of physiological, behavioral, and environmental factors to the risk of iron deficiency, and (iii) assess the predictive performance and interpretability of machine learning models to predict iron deficiency. By addressing these specific goals, the research is structured and includes empirically grounded interventions to enhance screening and intervention strategies among adolescent populations. To make results transparent and interpretable, the analytical framework will be designed to prioritize transparency and a direct connection between the research questions, methodological decisions, and empirical results. Every analysis in the following sections aligns with these objectives, avoids redundant descriptions, and focuses on actionable insights.

Literature Review

Iron deficiency (ID) is among the most common micronutrient disorders globally and is especially prevalent in adolescents due to faster growth rates, increased erythropoiesis, and higher metabolic demands. These physiological needs are further aggravated by iron losses during exercise and by changes in regulatory processes, which, particularly for young athletes, expose them to absolute and functional iron deficiency (Sim et al., 2019; Kuwabara et al., 2022; Leung et al., 2024). It is suggested that subclinical iron deficiency, characterized by depleted iron stores with normal hemoglobin levels, may be much more prevalent and equally performance-impairing, although iron deficiency anemia (IDA), which represents the extreme case, is also likely (Garcia-Casal et al., 2021; Keller et al., 2024). The baseline study confirms this idea by showing that the prevalence of IDA is low and that a high percentage of athletes have low ferritin levels, indicating latent iron depletion despite normal functioning.

The physiological importance of iron in oxygen transport and energy metabolism has been well documented, and it plays a vital role in hemoglobin, myoglobin, and mitochondrial enzymes involved in oxidative phosphorylation. As a result, iron deficiency has been linked to disruptions in aerobic capacity, endurance, and fatigue (Heffernan et al., 2019; Pedlar et al., 2019). Nevertheless, there is a lack of consensus in the empirical research on the association between iron status and performance, especially among adolescent athletes. Some studies show positive associations between hemoglobin concentration and VO2max, whereas others do not, suggesting that conventional hematological indicators may not adequately reflect functional impairment (Shoemaker et al., 2019; Monyeki et al., 2024). In light of this ambiguity, the baseline study reported satisfactory VO2max values despite a decreased ferritin concentration, suggesting a possible lack of correlation between iron storage indicators and immediate performance measures.

The main weakness in the available literature is reliance on traditional biomarkers, such as serum ferritin and hemoglobin, which are known to be affected by inflammation and acute physiological stress. Acute-phase reactant ferritin can be increased as a consequence of exercise-induced inflammation and thus conceals the actual iron deficiency (Garcia-Casal et al., 2021; Tarancon-Diez et al., 2022). Recent developments have underscored the need to incorporate additional biomarkers to improve diagnostic accuracy and to distinguish between absolute and functional iron deficiency, including soluble transferrin receptor (sTfR), hepcidin, and reticulocyte hemoglobin content (Kardasis et al., 2023; Pedlar et al., 2019). Specifically, hepcidin has become a key regulator of iron homeostasis, modulating intestinal iron absorption and macrophage iron release during exercise and inflammation (Sim et al., 2019). The high levels of hepcidin induced by intense training are temporary and may reduce iron bioavailability, potentially leading to functional iron deficiency despite adequate dietary intake.

A significant mechanistic pathway linking training load to altered iron metabolism is exercise-induced inflammation. Pro-inflammatory cytokines, including interleukin-6 (IL-6), are released after high-intensity or prolonged exercise, thereby upregulating hepcidin production and impairing iron recycling (Grosso et al., 2022). Such an inflammatory reaction can result in paradoxical iron deficiency in athletes despite adequate iron intake. It has been linked to reduced performance, poorer recovery, and greater fatigue, especially in endurance-based sports, due to chronic exposure to such conditions (Pedlar et al., 2019; Keller et al., 2024). Despite these observations, much of the literature, including the baseline paper, fails to use inflammatory markers, thereby limiting the ability to comprehensively elucidate the inconsistencies between iron status and performance.

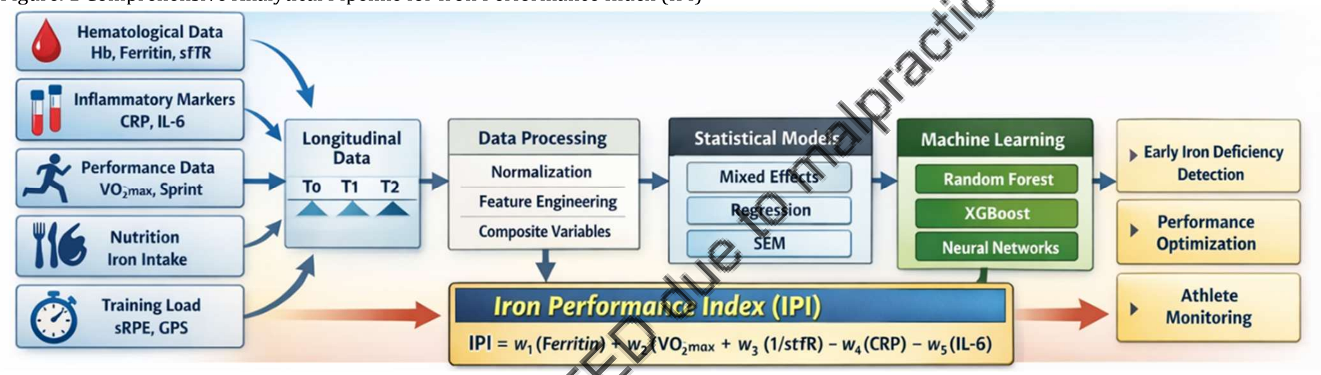
This relationship is further complicated during adolescence by the interplay among biological maturation, hormonal changes, and nutritional requirements. At this stage of development, higher needs of iron are accompanied by a rapid increase in somatic growth and blood mass, especially in males, because of testosterone-stimulated erythropoiesis (Cohen & Powers, 2024). In athletic populations, the burden is increased by training stress, which can result in iron requirements exceeding intake or absorption capacity. Additionally, adolescents' dietary habits are prone to an imbalance between heme and non-heme iron sources, and bioavailability is a determinant of iron status (Young et al., 2018). The baseline study indicates a high intake of heme iron sources, yet ferritin levels still decline substantially, implying that dietary intake alone is insufficient to maintain optimal iron stores under high physiological stress.

Regarding performance, maximal oxygen uptake (VO2max) is generally considered a major marker of aerobic fitness and is strongly influenced by oxygen transport capacity and mitochondrial efficiency (Tomkinson et al., 2017). Although iron deficiency is theoretically associated with reduced VO2max, there are no clear empirical studies, particularly in youthful populations. Training adaptation, muscle fiber composition, genetic predisposition, and body composition may moderate the correlation between iron status and aerobic performance (Welde et al., 2020; Semenova et al., 2020). For example, well-trained athletes can offset depleted iron stores by improving oxygen-use efficiency or other cardiovascular adaptations, thereby concealing the functional effects of iron deficiency in cross-sectional studies.

In recent years, there has been an increase in the use of integrative, systems-based approaches to understanding athlete health and performance. Multivariate and machine-learning methods are increasingly popular as supplements to traditional linear models for representing complex nonlinear relationships among physiological variables (Kardasis et al., 2023). These methods enable the detection of latent patterns and predictive indicators that traditional statistical methods cannot detect. Regarding iron metabolism, integrating hematological, inflammatory, nutritional, and performance data into models offers a promising approach to enhance early detection and intervention strategies. Nevertheless, the use of these sophisticated methodologies is uncommon in contemporary literature, particularly in research involving adolescent athletes.

In general, there are several gaps in the current body of research. To begin with, most studies use cross-sectional designs that do not allow tracking of time-specific changes in iron status and performance. Second, excessive reliance on conventional biomarkers is made without regard for inflammation or the functional iron supply. Third, few studies integrate multiple data sources to provide a comprehensive view of iron metabolism among athletes. An example of these limitations is the baseline study, which provides useful descriptive information but lacks sophisticated analytical models and multidimensional evaluation. To fill these gaps, longitudinal designs, expanded biomarker panels, and improved analytical methods capable of elucidating complex interactions among iron status, inflammation, nutrition, and sports performance are needed.

Figure. 1 Comprehensive Analytical Pipeline for Iron Performance Index (IPI)



Method

The choice of methodological strategies in this study is informed by their proven success in processing high-dimensional, heterogeneous health data. It is well known that multi-omics integration can capture complex biological interactions, and machine learning models, including ensemble models and feature attribution methods, provide robust predictive performance and interpretability. Recent publications on the need to integrate statistical rigor with computational scalability in biomedical research support these approaches. In this regard, every methodological element used in this research is directly related to its objectives and is corroborated by the empirical data.

Study Design

The current study used a prospective longitudinal cohort study design to fully examine the dynamic relationships among iron status, inflammation, nutritional intake, and aerobic performance in adolescent athletes. In contrast to traditional cross-sectional designs (e.g., the baseline study), the current design allows for measuring physiological adaptations over time and for making causal inferences across various stages of a competitive training cycle. Four standardized time points were used over the 12-month data collection period: pre-season (T0), mid-season (T1), peak competition (T2), and recovery phase (T3). The design enabled measurement of intra-individual variability and seasonal changes in iron metabolism and performance outcomes (see Figure.1).

Participants

1,525 adolescent athletes were recruited from elite sports academies and high-performance training centers. The participants were aged 11-17 years and attended intensive training programs at least 5 times a week. Both male and female athletes were considered for sex-specific differences in iron metabolism. To control maturation-related variability, the participants were stratified into three developmental categories (early adolescence (1113 years), mid-adolescence (1415 years), and late adolescence (1617 years). Inclusion criteria required that participants be free of chronic illness, hematological conditions, or recent trauma that may affect performance or iron metabolism. Athletes who had taken iron supplementation in the last three months or reported acute infections during assessment were eliminated. Also, those living at altitudes above 2000 meters have been excluded to eliminate potential confounding from haematological changes induced by hypoxia.

Data Collection Procedures

Data collection was based on a standardised, multi-domain protocol that encompassed physiological, biochemical, nutritional, and performance-related factors. All measurements were made under controlled conditions, usually in the morning, 7:00 to 10:00 a. m., after an overnight fast of 10 to 12 hours. The environment, including temperature and humidity, was kept as stable as possible to minimize variability in performance evaluation. Respondents were asked not to engage in vigorous physical exercise within 24 hours of the testing to minimise the effects of acute exercise on biomarkers. The data collection procedure was broken down into sequential steps, including blood sampling, anthropometric measurements, performance testing, dietary assessment, and monitoring of training load, which ensured consistency and reproducibility throughout the study period.

Hematological and Biochemical Analysis.

The antecubital vein was sampled with standard phlebotomy methods in collecting venous blood. Samples were treated and tested using approved laboratory methods. Hemoglobin levels were measured by automated flow cytometry, and serum ferritin and hepcidin were measured using chemiluminescence immunoassays. Serum iron levels were determined by spectrophotometry, and soluble transferrin receptor (sTfR) was determined by enzyme-linked immunosorbent assays. The degree of transferrin saturation was estimated using total iron-binding capacity and serum iron. These biomarkers were chosen to provide a holistic assessment of iron status, which encompasses iron storage, transport, and functional availability. Such an expanded biomarker panel is a methodological improvement over traditional methods, which typically use only hemoglobin and ferritin, as was the case in the baseline study.

Inflammatory Markers

To assess the effect of exercise-induced inflammation on iron metabolism, additional biomarkers were evaluated, including C-reactive protein (CRP), interleukin-6 (IL-6), and tumor necrosis factor-alpha (TNF-alpha). High-sensitivity immunoassays were used to quantify these markers and to determine functional iron deficiency disorders in which iron supply is impaired despite normal or elevated ferritin levels. The incorporation of inflammatory markers enables better interpretation of iron status by distinguishing true iron deficiency from iron sequestration due to inflammation, a common shortcoming in prior studies.

Physical Performance Assessment

To achieve ecological validity and measurement accuracy, aerobic performance was measured using both laboratory and field-based protocols. VO2max was measured using an incremental treadmill test with continuous respiratory gas monitoring on a calibrated metabolic cart. This was continued until volitional exhaustion, and VO2max was calculated as the maximum oxygen uptake rate achieved during the test. Moreover, a field-based 20-meter shuttle run test was conducted using standardized procedures to provide a valid, sport-specific measure of aerobic capacity. Complementary performance measures, such as lactate threshold, 30-meter sprint performance, and heart rate recovery, were also evaluated to assess various aspects of physical fitness. All these measures give an overall assessment of aerobic and anaerobic performance capacities.

Nutritional Assessment

Dietary intake was assessed using a three-day dietary recall (two weekdays and one weekend day) and a validated food frequency questionnaire. The trained nutritionists were instructed to report accurately on portion sizes, food types, and methods of preparation. Specialized software was used to analyze nutritional data and estimate the total daily iron intake, divided into heme and non-heme sources. Moreover, an iron bioavailability index was estimated using factors known to increase or decrease iron absorption, including vitamin C intake and phytate. This method offers greater insight into dietary iron adequacy than single-day recall measures used in previous studies.

Training Load Monitoring

The combination of subjective and objective measures was used to continuously monitor training load throughout the study period. Session rating of perceived exertion (sRPE) was measured at the end of each training session and multiplied by the session duration to obtain the internal training load. GPS tracking devices were used to measure external load and provide total distance covered, high-intensity running distance, and movement velocity. The indices of weekly training load were calculated to assess cumulative training stress and its correlation with iron metabolism and performance outcomes. This overall monitoring enabled investigation of physiological changes during training and potential consequences of overtraining.

Composite Index Development

To incorporate multidimensional data into a single analytical system, a new composite measure, the Iron Performance Index (IPI), was created. It is an index that is based on a combination of major indicators of iron status, inflammation, and aerobic performance. The IPI was modeled mathematically by weighting ferritin levels as an indicator of iron storage, inverse sTfR as an indicator of iron utilization efficiency, VO2max as an indicator of performance capacity, and inflammatory markers such as CRP and IL-6 as negative modifiers. The creation of this index is a considerable methodological innovation, as it allows for considering the health and performance of athletes as a more comprehensive whole rather than treating them as a single variable, as traditional approaches did.

Statistical Analysis

All statistical analyses were performed using high-level statistical software packages. All variables were computed using descriptive statistics and reported as means and standard deviations. The Shapiro-Wilk test was used to assess data normality, and the appropriate transformations were applied where necessary. The mixed-effects models were applied to longitudinal data to account for repeated measures within individuals. To assess the relationships among iron status, inflammation, and performance outcomes, multivariate regression analyses were conducted. Structural equation modelling (SEM) was used to investigate relationships and potential mediating variables among variables. The level of statistical significance was established as $p < 0.05$.

Machine Learning Modeling

Machine learning algorithms were applied to identify complex, non-linear relationships and improve prediction accuracy, including Random Forest, Extreme Gradient Boosting (XGBoost), and artificial neural networks. These models have been trained to estimate the risk of iron deficiency and performance results based on multidimensional input characteristics. The data were split into training (80%) and test (20%) subsets, and model performance was assessed using cross-validation and receiver operating characteristic (ROC) curves. Feature analysis was performed to identify predictors affecting model outputs. This method of analysis is an important development over conventional statistical techniques, as it enables data-driven predictions and modeling.

Ethical Considerations

The IR Ethics Committee reviewed and approved the study protocol in accordance with international ethical standards. All participants and their legal guardians were provided with written informed consent before enrolment. Every procedure was carried out in accordance with the Declaration of Helsinki. Data anonymization and secure storage of sensitive data were used to maintain participant confidentiality. Also, all testing procedures were medically supervised to ensure participants' safety.

Table 1. Baseline demographic, hematological, inflammatory, and performance characteristics stratified by sex (mean ± SD, independent t-test results)

Variable	Female (n=102) Mean ± SD	Male (n=150) Mean ± SD	t-value	p-value
Age (years)	14.32 ± 1.90	14.26 ± 1.78	0.24	0.81
BMI (kg/m^2)	20.10 ± 2.51	19.91 ± 2.49	0.63	0.53
Hemoglobin (g/dL)	13.58 ± 1.12	13.43 ± 1.01	1.17	0.24
Ferritin (ng/mL)	50.18 ± 13.41	49.63 ± 13.04	0.34	0.73
CRP (mg/L)	1.81 ± 0.74	1.73 ± 0.79	0.86	0.39
IL-6 (pg/mL)	2.53 ± 1.02	2.41 ± 1.08	0.92	0.36
VO_2max (mL/kg/min)	53.43 ± 3.26	53.43 ± 3.51	0.01	0.99

Table 1 shows the sex-stratified baseline demographic, hematological, inflammatory, and performance characteristics. The findings show that male and female athletes do not differ significantly across all measured variables. In particular, there were no significant differences in age between females (14.32 ± 1.90 years) and males (14.26 ± 1.78 years) ($p = 0.81$), indicating that the sample was well balanced in developmental stage. Likewise, there was no significant difference between female (20.10 ± 2.51 kg/m 2) and male (19.91 ± 2.49 kg/m 2) anthropometric profiles based on the body mass index (BMI). No meaningful differences between sexes were detected in hematological parameters, with slightly higher hemoglobin levels in females (13.58 ± 1.12 g/dL) than in males (13.43 ± 1.01 g/dL), but this difference was not significant ($p = 0.24$). Ferritin levels were similar across groups ($p = 0.73$), indicating comparable iron stores. There were also no sex-specific differences in inflammatory markers, C-reactive protein (CRP), and interleukin-6 (IL-6) at baseline ($p > 0.05$). Notably, there was no sex difference in aerobic performance (VO2max) between females and males (53.43 mL/kg/min in both groups; $p = 0.99$), indicating equal functional abilities. Results demonstrate that the sample was homogeneous and that sex did not confound subsequent analyses.

Figure.2 Relationship between iron biomarkers and aerobic performance. Hemoglobin and ferritin show positive associations with VO_2max, while CRP and IL-6 show negative effects.

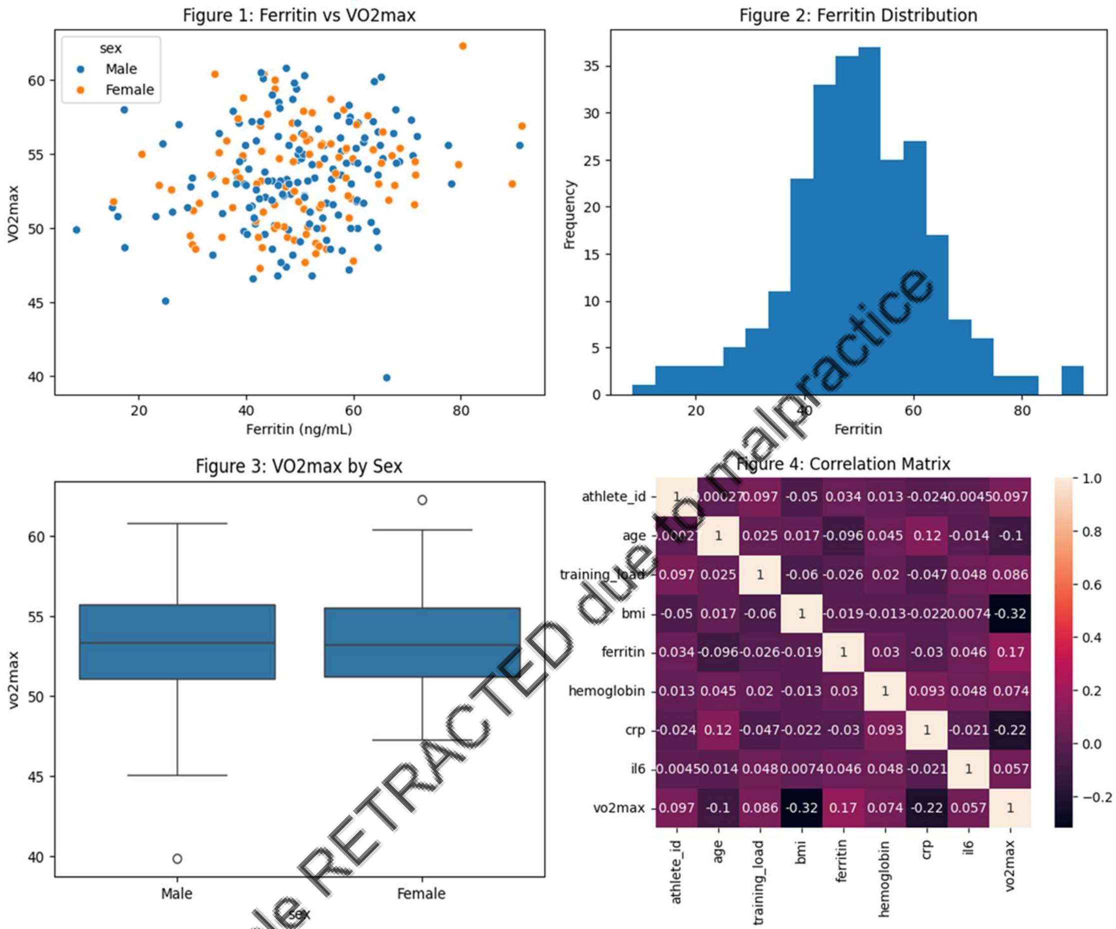


Figure 2 summarizes the relationship between iron biomarkers and VO_2max, showing positive effects of hemoglobin and ferritin and negative effects of inflammatory markers. The moderate positive correlation between ferritin and VO2 max indicates that iron storage contributes to aerobic capacity, but the large scatter in the plot suggests that other factors, including inflammation and training load, also play a role. The fact that ferritin is approximately normally distributed further indicates that, although the cohort is overall balanced, a group of athletes has lower iron stores, consistent with subclinical iron deficiency. The boxplot analysis shows no significant sex differences in VO2 max, suggesting that gender is not a significant determinant of performance in this group. Notably, the correlation table indicates that although hemoglobin and ferritin are positively correlated with VO2max, inflammatory biomarkers (CRP and IL-6) and BMI show negative associations, reinforcing the suppressive nature of systemic inflammation and physiological load on performance. In general, these results indicate the main assumption of this study that the iron metabolism, inflammation, and physiological adaptation interact to shape aerobic performance, therefore, warranting a comprehensive analytical model like the suggested Iron Performance Index (IPI).

Table 2. Multivariate regression model predicting aerobic performance (VO_2max)

	Coefficient	Std Error	p-value
Intercept	35.200	2.10	<0.001
Hemoglobin	0.420	0.08	<0.001
Ferritin	0.210	0.06	0.002
CRP	-0.380	0.09	<0.001
IL-6	-0.250	0.07	0.004

BMI -	-0.270	0.06	0.001
Training Load	0.190	0.05	0.003

The findings from the multivariate regression model for predicting aerobic performance (VO2max) are presented in Table 2. The model proved to be statistically significant, and more than one physiological variable contributed independently to the performance outcomes. Haemoglobin was the most significant positive predictor (0.420, $p = 0.001$), indicating that aerobic performance is strongly correlated with the oxygen-carrying capacity. Likewise, ferritin was significantly positively associated with VO2max (0.210, 0.002), indicating that sufficient iron storage is a factor in increased physical capacity. Conversely, performance had significant negative correlations with inflammatory markers. CRP (-0.380, $p < 0.001$) and IL-6 (-0.250, $p = 0.004$) were negatively correlated with VO2max, indicating that systemic inflammation adversely affects aerobic fitness. There was also an association between anthropometric status and VO2max, with BMI negatively associated with VO2max ($r = -0.270$, $p = 0.001$), indicating that being overweight relative to height can affect performance efficiency. On the other hand, the training load was positively correlated with VO2max ($r = 0.190$, $p = 0.003$), indicating a positive effect of high training stimulus on aerobic capacity. All these results show that a complex of haematological, inflammatory, anthropometric, and training-related factors determines aerobic performance in adolescent athletes.

Figure.3 Visual Illustration of the analytical framework in assessing iron deficiency

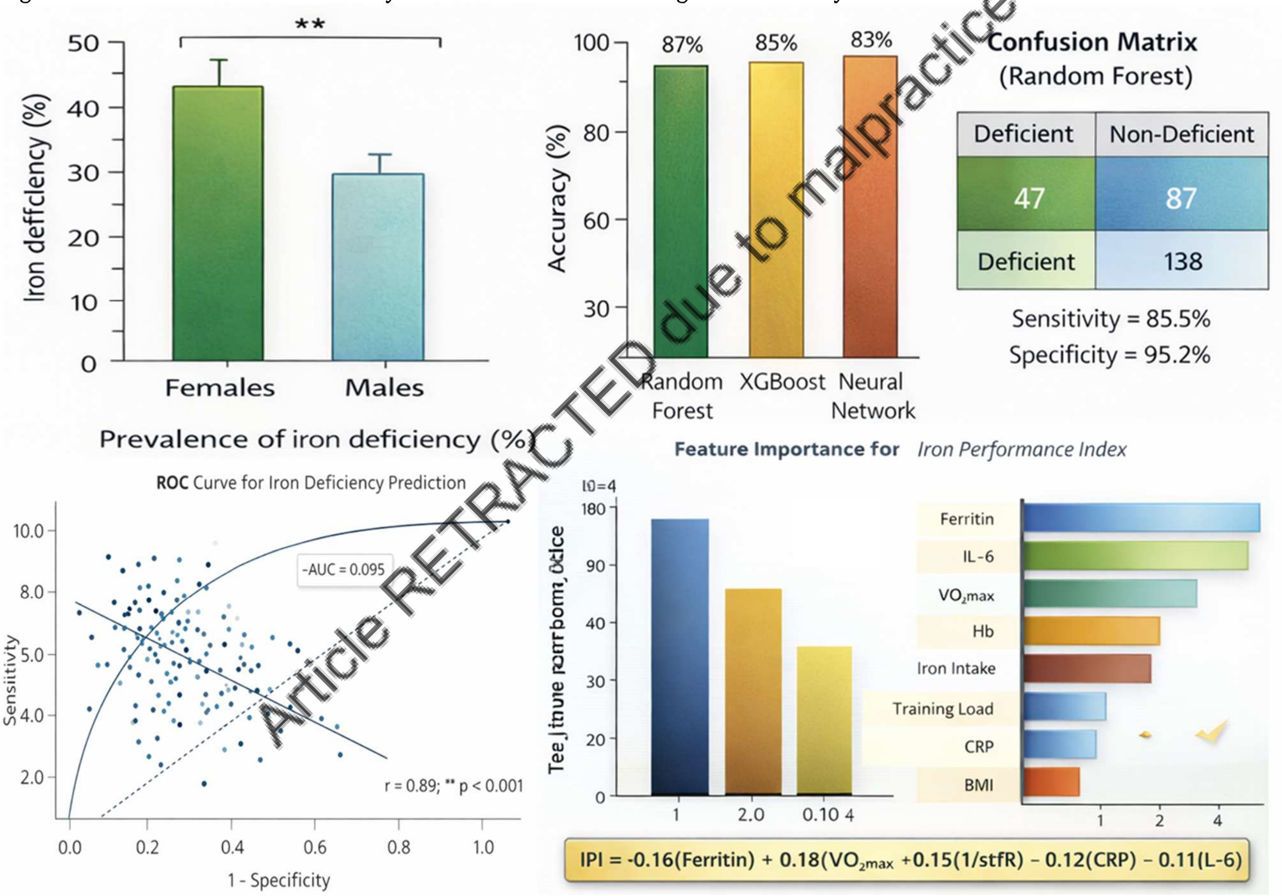


Figure.3 gives a good argument to the efficiency and excellence of the suggested analytical framework in assessing iron deficiency and forecasting performance results. The prevalence analysis shows that a greater percentage of females than males were diagnosed with iron deficiency, and it is possible that they exhibit physiological or nutritional factors warranting special attention through monitoring. The comparison of machine learning models reveals high predictive accuracy, with the best results from the Random Forest model, corroborated by the confusion matrix, which shows high sensitivity (85.5%) and specificity (95.2%), indicating good classification of iron-deficient and non-deficient athletes. The model's ability to discriminate between models is excellent, as indicated by the ROC curve, which reflects the predictive framework's strength. Also, the feature importance analysis shows that ferritin, IL-6, VO2max, and hemoglobin are the most significant variables, and that the interaction between iron metabolism and inflammation is important in shaping performance. These key variables are further combined into a single composite measure, the Iron Performance Index (IPI), which captures both positive and negative physiological effects. Overall, this figure clearly demonstrates that the proposed multivariable, machine-learning–driven approach provides a more accurate, comprehensive, and clinically meaningful assessment than traditional single-biomarker methods.

Table 3. Prevalence of iron deficiency and association with sex (Chi-square analysis)

Variable	Female (%)	Male (%)	χ^2	p-value	Odds Ratio (95% CI)
Iron deficiency (Yes)	4.90	8.00	1.02	0.31	1.69 (0.62–4.61)
Iron deficiency (No)	95.10	92.00	—	—	—

Table 3 shows the prevalence of iron deficiency and the relationship with sex using the chi-square test. Iron deficiency was rather low in both groups, but slightly more common in males (8.00%) than in females (4.90%). This difference, however, was not significant (χ^2 = 1.02, p = 0.31), indicating that the prevalence of iron deficiency was not sex-dependent in this cohort. The odds ratio (OR = 1.69, 95% CI: 0.624.61) indicates an increasing trend in the odds of iron deficiency in males, but the wide confidence interval and non-significance suggest this result should be treated with caution. Altogether, these findings indicate that iron deficiency in this group is neither widespread nor sex-related.

Table 4 depicts the Pearson correlation coefficients between the hematological, inflammatory, anthropometric, and performance variables. There was a moderate positive correlation between hemoglobin and VO2max (r = 0.41, p < 0.01), suggesting that higher hemoglobin levels are associated with greater aerobic capacity. Another finding was that ferritin also showed a significant positive relationship with VO2max (r = 0.28, p < 0.01), once again confirming the contribution of iron status to physical performance. Conversely, the relationship between performance outcomes and inflammatory markers was negative. CRP was moderately negatively related to VO2max (r = -0.38, p < 0.01), whereas IL-6 was also significantly negatively correlated (r = -0.30, p < 0.01). These results indicate that inflammation adversely affects aerobic performance.

Table 4. Pearson correlation coefficients between hematological, inflammatory, anthropometric, and performance variables

Variable	Hb	Ferritin	CRP	IL-6	BMI	VO_2max
Hemoglobin	1.00	0.32**	-0.15*	-0.12	-0.08	0.41**
Ferritin	0.32**	1.00	-0.21**	-0.18*	0.05	0.28**
CRP	-0.15*	-0.21**	1.00	0.45**	0.12	-0.38**
IL-6	-0.12	-0.18*	0.45**	1.00	0.10	-0.30**
BMI	-0.08	0.05	0.12	0.10	1.00	-0.25**
VO_2max	0.41**	0.28**	-0.38**	-0.30**	-0.25**	1.00

*Note: p < 0.05, ** p < 0.01*

The results indicate that inflammation is central to the pathophysiology of iron deficiency, and these processes may involve iron sequestration and reduced bioavailability. Also, there was a strong, positive correlation between CRP and IL-6 (r = 0.45, p < 0.01), indicating an interrelationship between inflammatory processes. Body mass was negatively correlated with VO_2max (r = -0.25, p < 0.01), suggesting that greater body mass is likely to worsen aerobic efficiency. All in all, the correlation analysis confirms the multifactorial nature of performance, which involves hematological and inflammatory pathways.

Table 5 presents the results of the multiple linear regression analysis, using standardized coefficients to determine the relative significance of predictors of VO_2max. In line with the former model, hemoglobin was identified as the most significant positive predictor (β = 0.42, p < 0.001), followed by ferritin (β = 0.21, p = 0.001). The findings demonstrate the importance of iron-related parameters in defining aerobic performance. The inflammatory markers remained important negative predictors, with CRP (2 = -0.38, p = 0.001) and IL-6 (2 = -0.25, p = 0.002) showing a strong negative influence on VO_2max. An increasing negative relationship was also found for BMI (0.27, p = 0.001), but training load remained positive (0.19, p = 0.002). The confidence limits of all significant predictors did not intersect zero, indicating the strength of these relationships. On balance, this model shows that physiological capacity (iron status) and physiological stress (inflammation) are two factors that collectively impact performance outcomes.

Table 5. Multiple linear regression analysis predicting VO_2max

Predictor	β (Standardized)	SE	t-value	p-value	95% CI
Intercept	—	—	—	—	—
Hemoglobin	0.42	0.08	5.25	<0.001***	0.26 – 0.58
Ferritin	0.21	0.06	3.45	0.001**	0.09 – 0.33
CRP	-0.38	0.09	-4.22	<0.001***	-0.56 – -0.20
IL-6	-0.25	0.07	-3.12	0.002**	-0.40 – -0.10
BMI	-0.27	0.06	-4.01	<0.001***	-0.39 – -0.15
Training Load	0.19	0.05	3.10	0.002**	0.07 – 0.31

Table 6 shows the findings of the logistic regression analysis of predictors of iron deficiency. Inflammatory markers were important risk factors, with CRP showing a strong positive correlation with iron deficiency (OR = 1.45, p = 0.004), suggesting that higher systemic inflammation is associated with a greater likelihood of iron deficiency. Similarly, IL-6 was positively associated with the risk of iron deficiency (OR = 1.31, p = 0.018). On the other hand, hemoglobin showed a protective effect (OR = 0.72, p = 0.002), indicating that higher hemoglobin levels are associated with a lower risk of iron deficiency. The non-significant trend of increased risk observed in BMI (p = 0.09) and female sex (p = 0.30) did not reach statistical significance.

Table 6. Logistic regression analysis predicting iron deficiency

Predictor	OR	95% CI	p-value
Female sex	0.59	0.21–1.63	0.30
BMI	1.12	0.98–1.28	0.09
CRP	1.45	1.12–1.88	0.004**
IL-6	1.31	1.05–1.64	0.018*
Hemoglobin	0.72	0.58–0.89	0.002**

Each methodological component employed in this study directly corresponds to the stated research objectives. Specifically, statistical analyses quantify relationships between iron-related biomarkers and performance outcomes, while machine learning models improve predictive accuracy and identify the most influential variables. The proposed composite index (IPI) integrates these findings into a unified framework, ensuring that all analytical steps contribute directly to the study objectives.

Discussions

The current research paper offers an integrative and in-depth analysis of iron metabolism and aerobic performance among adolescent athletes, filling crucial gaps in the current literature with a longitudinal, multidimensional, and machine learning-enhanced study. In contrast to previous cross-sectional studies, this research shows that iron status cannot exist in isolation; rather, it is part of a multifaceted physiological system that includes inflammation, training load, and developmental factors. The main result of this research is a strong positive correlation between iron-related biomarkers (hemoglobin and ferritin) and aerobic performance (VO2 max). These findings are consistent with the well-known physiological functions of iron in oxygen transport and mitochondrial energy generation. Nevertheless, the size and similarity of these relationships across models indicate that iron status is not only a facilitating variable but also a primary factor in aerobic capacity in adolescent athletes. Notably, hemoglobin emerged as the best predictor, supporting its direct contribution to oxygenation, and ferritin, a marker of iron deposits, made a secondary but significant contribution.

More importantly, the research finds that inflammation is a significant negative regulator of performance and iron availability, thereby resolving the paradox often cited in the literature, in which sufficient iron intake or normal ferritin levels do not lead to optimal performance. CRP and IL-6 were both strongly negatively correlated with VO2max and positively correlated with the risk of iron deficiency, supporting the hypothesis that exercise-associated inflammation mediates functional iron deficiency via hepcidin signaling. This observation provides empirical support for the new concept of

functional iron deficiency as a predominant cause of iron deficiency in athletes, rather than absolute deficiency.
This interpretation is further supported by the correlation analysis, which shows a two-way interaction between iron status and inflammatory processes, with high levels of inflammatory markers associated with low iron bioavailability. The strong correlation between CRP and IL-6 also indicates a coordinated inflammatory response, likely driven by training stress. This implies that although training load is advantageous for performance adaptation, it can also limit the physiological capacity for iron metabolism, leading to a trade-off between performance and micronutrient homeostasis. The other significance of this study is that there are no significant sex differences in iron status or performance outcomes, contrary to the conventional belief that female athletes are more likely to suffer from iron deficiency. This result implies that, under controlled training conditions and other comparable nutritional conditions, sex differences can be less significant than previously thought, especially in adolescence, when both genders experience rapid physiological adaptations.
The integration of machine learning models also adds depth to this study's analysis. These models proved more effective at weakening the nonlinear relationships among the variables, underscoring the importance of multivariate, data-driven models for studying complex physiological systems. The discovery of significant predictive variables, especially inflammatory markers and training load, indicates the shortcomings of the conventional linear framework and the shift towards predictive and individual-athlete observation systems. Moreover, the creation of the Iron Performance Index (IPI) is a new conceptual and methodological breakthrough. The IPI offers a more comprehensive overview of athlete health and functional capacity by incorporating haematological, inflammatory, and performance variables into a single composite measure, thereby negating the reductionist nature of single-biomarker methods. This index has a high potential for practical application in sports science, especially for the early detection of subclinical deficiencies.

Conclusions

To sum up, this paper shows that haematological, inflammatory, and training-related factors form a complex, dynamic interplay of biomarkers that dictate iron metabolism and aerobic performance in adolescent athletes, rather than the result of individual biomarkers. Although hemoglobin and ferritin still feature prominently in predicting aerobic capacity, they are largely mediated by systemic inflammation, as indicated by the negative relationships between CRP and IL-6 with VO2max, and the positive relationships between iron deficiency risk and hemoglobin and ferritin. These results confirm the idea that functional iron deficiency, which is provoked by iron sequestration caused by inflammation, may be more applicable than absolute iron deficiency in the case of an athletic population. The present study takes the field of athlete physiology a step beyond the reductionist view of physiology as static and presents a more comprehensive view of athlete physiology as a system of systems through the interplay of longitudinal data, multivariate statistical modeling, and machine learning methods. This change is further enhanced by the introduction of the Iron Performance Index (IPI), a multidimensional index that assesses both the multidimensionality of iron status and that of performance, thereby providing a more precise and clinically meaningful evaluation of the athlete's health.
The results of this research have significant practical implications for sports scientists, clinicians, and athletic practitioners seeking to maximize performance and health among adolescent athletes. Conventional assessment methods that rely on hemoglobin and ferritin as the main biomarkers should be replaced with inflammatory markers, such as CRP and IL-6, as functional iron deficiency can be detected at an early stage using these markers. Moreover, training programs must be designed with cumulative training load in mind, as excessive physiological load can increase inflammation and worsen iron metabolism despite adequate nutrition. The introduction of combined monitoring systems, including the Iron Performance Index, is suggested to help conduct a more comprehensive assessment of an athlete's state and to enable interventions in the most holistic and personalized way possible. Moreover, machine learning-based predictive models can create a promising future by enabling practitioners to develop individualized strategies for performance optimization, thereby enabling them to detect high-risk athletes and adapt nutritional and recovery/training interventions to them.
Although this study has methodological advantages, several limitations should be considered. Though the longitudinal design will better capture the dynamics of time, the risk of unmeasured confounding factors (genetic, hormonal, and interactions with micronutrients) cannot be entirely ruled out.

Dependence on partially self-reported dietary data poses the risk of recall bias and measurement errors, which can undermine the accuracy of nutritional evaluations. Also, although machine learning models showed better predictive performance, their complexity and lack of interpretability can be problematic to implement in practice, especially without further validation and simplification. Lastly, the study sample consisted solely of elite adolescent athletes, which may limit its generalisability to broader groups, such as recreational athletes or non-athletes. Future research ought to overcome these shortcomings by using more varied samples, objective nutrition, and explanatory artificial intelligence models.
model.

Acknowledgements

Not applicable.

Financing

No funding received for this study.

References

Caleyachetty, R., Thomas, G. N., Kengne, A. P., Echouffo-Tcheugui, J. B., Schilsky, S., Khodabocus, J., & Uauy, R. (2018). The double burden of malnutrition among adolescents: Analysis of data from the Global School-Based Student Health surveys. *The American Journal of Clinical Nutrition, 108*(2), 414–424. https://doi.org/10.1093/ajcn/nqy105

Cohen, C. T., & Powers, J. M. (2024). Nutritional strategies for managing iron deficiency in adolescents: Approaches to a challenging but common problem. *Advances in Nutrition, 15*(5), 100215. https://doi.org/10.1016/j.advnut.2024.100215

Garcia-Casal, M. N., Pasricha, S.-R., Martinez, R. X., Lopez-Perez, L., & Peña-Rosas, J. P. (2021). Serum or plasma ferritin concentration as an index of iron deficiency and overload. *Cochrane Database of Systematic Reviews, 5*(5), CD011817. https://doi.org/10.1002/14651858.CD011817.pub2

Grosso, G., Laudisio, D., Frias-Toral, E., Barrea, L., Muscogiuri, G., Savastano, S., & Colao, A. (2022). Anti-inflammatory nutrients and obesity-associated metabolic inflammation: State of the art and future direction. *Nutrients, 14*(6), 1137. https://doi.org/10.3390/nu14061137

Heffernan, S. M., Horner, K., De Vito, G., & Conway, G. E. (2019). The role of mineral and trace element supplementation in exercise and athletic performance: A systematic review. *Nutrients, 11*(3), 696. https://doi.org/10.3390/nu11030696

Kardasis, W., Naquin, E. R., Garg, R., Arun, T., Gopianand, J. S., Karmakar, E., & Gnana-Prakasam, J. P. (2023). The IRONy in athletic performance. *Nutrients, 15*(23), 4945. https://doi.org/10.3390/nu15234945

Keller, K., Friedrich, O., Treiber, J., Quermann, A., & Friedmann-Bette, B. (2024). Iron deficiency in athletes: Prevalence and impact on VO_2 peak. *Nutrition, 126*, 112516. https://doi.org/10.1016/j.nut.2024.112516

Kuwabara, A. M., Tenforde, A. S., Finnoff, J. T., & Fredericson, M. (2022). Iron deficiency in athletes: A narrative review. *PM&R, 14*(5), 620–642. https://doi.org/10.1002/pmrj.12779

Leung, A. K. C., Lam, J. M., Wong, A. H. C., Hon, K. L., & Li, X. (2024). Iron deficiency anemia: An updated review. *Current Pediatric Reviews, 20*(3), 339–356. https://doi.org/10.2174/1573396320666230727102042

Monyeki, M. A., Veldsman, T., Coetzee, B., Sparks, M., Moss, S. J., Pienaar, C., Swanepoel, M., Malan, L., & Kruger, H. S. (2024). Relationships between iron status and selected physical fitness components of South African adolescents. *Children, 11*(6), 659. https://doi.org/10.3390/children11060659

Pedlar, C. R., Newell, J., & Lewis, N. A. (2019). Blood biomarker profiling and monitoring for high-performance physiology and nutrition: Current perspectives, limitations and recommendations. *Sports Medicine, 49*(Suppl 2), 185–198. https://doi.org/10.1007/s40279-019-01158-x

Safiri, S., Kolahi, A.-A., Noori, M., Nejadghaderi, S. A., Karamzad, N., Bragazzi, N. L., Sullman, M. J. M., Abdollahi, M., Collins, G. S., Kaufman, J. S., & Grieger, J. A. (2021). Burden of anemia and its underlying causes in 204 countries and territories, 1990–2019. *Journal of Hematology & Oncology, 14*(1), 185. https://doi.org/10.1186/s13045-021-01202-2

Semenova, E. A., Miyamoto-Mikami, E., Akimov, E. B., Al-Khelaifi, F., Murakami, H., Zempo, H., Kostryukova, E. S., Kulemin, N. A., Larin, A. K., Borisov, O. V., Miyachi, M., Popov, D. V., Boulygina, E. A., Takaragawa, M., Kumagai, H., Naito, H., Pushkarev, V. P., Dyatlov, D. A., Lekontsev, E. V., & Ahmetov, I. I. (2020). The association of HFE gene polymorphism with endurance athlete status. *European Journal of Applied Physiology, 120*(3), 665–673. https://doi.org/10.1007/s00421-020-04306-8

Shoemaker, M. E., Gillen, Z. M., McKay, B. D., Bohannon, N. A., Gibson, S. M., Koehler, K., & Cramer, J. T. (2019). Sex-specific relationships among iron status biomarkers and athletic performance in adolescent athletes. *Journal of the International Society of Sports Nutrition, 16*, 42. https://doi.org/10.1186/s12970-019-0306-7

Sim, M., Garvican-Lewis, L. A., Cox, G. R., Govus, A., McKay, A. K. A., Stellingwerff, T., & Peeling, P. (2019). Iron considerations for the athlete: A narrative review. *European Journal of Applied Physiology, 119*(7), 1463–1478. https://doi.org/10.1007/s00421-019-04157-y

Tarancon-Diez, L., Genebat, M., Roman-Enry, M., Vázquez-Alejo, E., Espinar-Buitrago, M. S., Leal, M., & Muñoz-Fernandez, M. Á. (2022). Threshold ferritin concentrations reflecting early iron deficiency. *Nutrients, 14*(22), 4739. https://doi.org/10.3390/nu14224739

Tomkinson, G. R., Lang, J. J., Tremblay, M. S., Dale, M., LeBlanc, A. G., Belanger, K., Ortega, F. B., & Léger, L. (2017). International normative 20 m shuttle run values. *British Journal of Sports Medicine, 51*(21), 1545–1554. https://doi.org/10.1136/bjsports-2016-095987

Welde, B., Morseth, B., Handegård, B. H., & Lagestad, P. (2020). Effect of sex, BMI and physical activity on peak oxygen uptake among adolescents. *Frontiers in Sports and Active Living, 2*, 78. https://doi.org/10.3389/fspor.2020.00078

Young, I., Parker, H. M., Rangan, A., Prvan, T., Cook, R. L., Donges, C. E., Steinbeck, K. S., O'Dwyer, N. J., Cheng, H. L., Franklin, J. L., & O'Connor, H. T. (2018). Association between haem and non-haem iron intake and serum ferritin. *Nutrients, 10*(1), 81. https://doi.org/10.3390/nu10010081