Establishing Reference and Optimal Curves for Hemoglobin and Ferritin: Methodological and Computational Frameworks to Evaluate and Synthesize Existing and Emerging Evidence

VID BIJELIĆ BSc, MSc

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Doctor of Philosophy in Epidemiology

School of Epidemiology and Public Health
Faculty of Medicine
University of Ottawa

To Marijana

Supervisors

Primary supervisor: Dr. Beth K. Potter, University of Ottawa

Co-supervisor: Dr. Jemila S. Hamid, University of Ottawa

Thesis advisory committee:

Dr. Patricia C. Parkin, Hospital for Sick Children Research Institute

Dr. Mira Liebman, University of Ottawa

Dr. Franco Momoli, University of Ottawa

PREFACE TO THESIS

Funding
This work was supported by a tuition scholarship from School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa (2020-2024) and a Natural Sciences and Engineering Research Council of Canada, Government of Canada – JSH research grant (number: RGPIN-2018-06693) (2020-2024).

Ethical Considerations
Ethics approval for Projects II and III of this thesis (requiring analysis of pediatric data from a Canadian cohort) was obtained from the Research Ethics Boards at Clinical Trials Ontario Board of Record SickKids Hospital (#2063) and the University of Ottawa Research Ethics Board (#H-10-24-10933). The author of this dissertation, Vid Bijelić, is a student and an employee at SickKids and CHEO Hospital.

Author Contributions
With the supervision of my thesis supervisors Dr. Jemila Hamid and Dr. Beth Potter and thesis advisory committee members, I, Vid Bijelić (the PhD Candidate) am the guarantor of the thesis projects. The Thesis Advisory Committee (Dr. Patricia C. Parkin, Dr. Mira Liebman, Dr. Franco Momoli) was selected for their clinical, pediatrics, hematology, and epidemiology expertise. Vid Bijelić is the primary and the first, author of all papers that were published or submitted for publications. I conceptualized all the projects, developed thesis protocol, developed novel methods and/or chosen appropriate statistical methods, conducted data analysis, interpreted the results, drafted all the manuscript. I am the sole author of this thesis. Supervisors and thesis advisory committee members contributed to the critical review of the thesis protocol, methods and the findings and provided critical revisions of the all the manuscripts and the thesis. Additional coauthors were invited to collaborate on selected manuscripts as appropriate.

Dissertation abstract

Introduction: Reference intervals (RIs) provide benchmarks for interpreting laboratory test results; however, inconsistencies in their development complicate clinical decision-making, especially in pediatrics where biomarker levels change rapidly with age and development. My dissertation synthesized existing evidence on pediatric RIs and reference curves (RCs) for hemoglobin and ferritin and established RCs and optimal curves (OCs) for these biomarkers using new methodological and computational frameworks. Although developed for hemoglobin and ferritin, these frameworks can be applied to other pediatric biomarkers and tailored to other populations.

Methods: I conducted two systematic reviews with meta-analyses to understand existing evidence and methodological challenges in generating pediatric hemoglobin and ferritin RIs and RCs. I evaluated heterogeneity using forest plots, heatmaps, web-based visualization tools, and the I^2 statistic. I also proposed a standardized age partitioning approach to enable quantitative synthesis. I then analyzed data from the TARGet Kids! cohort study of healthy Canadian children ages 2-weeks to 10-years. I developed RCs for pediatric hemoglobin and ferritin, overcoming excessive partitions and small sample sizes in early childhood. I also introduced novel pediatric OCs derived from a sub-sample of participants meeting predefined optimality criteria. Finally, I created web-based tools for both the systematic reviews and cohort analyses, to explore heterogeneity, visualize data, and aid in interpretation.

Findings: Both systematic reviews showed substantial heterogeneity across studies due to differing age intervals, population characteristics, and analyzer types, with limited data available for very young children. Lower limits from many published RIs differed from published World Health Organization (WHO) thresholds. Among participants in our full cohort analysis and those

meeting predefined optimality criteria, the proportion of pre-adolescents who would be classified as iron deficient according to the WHO thresholds and the American Society of Hematology (ASH) thresholds varied substantially with age.

Conclusion: RCs and OCs provide essential benchmarks for interpreting results, evaluating thresholds, and assessing population iron status. This work identified key methodological limitations, addressed several of these, and established frameworks for synthesizing evidence and developing pediatric RCs and OCs. Future studies should evaluate the clinical value of the tools developed and extend these approaches to other key laboratory biomarkers.

Acknowledgements

This work would not see the daylight and come to an end without a dedicated support team. I am profoundly grateful to each of you.

I would like to say big thank you to my supervisors, Professor Dr. Jemila Hamid and Professor Dr. Beth Potter. Professor Hamid whose expertise, mentorship, and endless encouragement guided me to the finishing line. Thank you immeasurably for hour countless conversations at CHEO and giving me vast possibilities when I didn't think about them. Professor Dr. Potter endless support and guidance was unwavering not only in this work, but the first time I entered the classroom for my first EPI5240 class. I have been blessed for having you as my supervisors. I would also like to thank Professors Dr. Patricia Parkin, Dr. Mira Liebman, and Dr. Franco Momoli my thesis advisory committee members, without their clinical and epidemiological expertise this thesis would not be possible. My sincere thank you goes to Dr. Parkin for her immense guidance, expertise, and collaboration at our TAC meetings. I have been very lucky to have you all on my team. I will miss our meetings.

Josh and Michael, thank you for your laughs and help with numerous data extractions.

My mom, Stana, for following my journey from Republika Srpska. Having countless telephone conversations. Thank you.

Special thank you to my friends Darko and Jelena Topić for their unreserved support throughout my journey.

Most importantly, I am grateful to my wife Marijana on her unconditional support and infinite patience. The journey would be impossible without you.

Table of Contents

List of Tables

List of Figures

CHAPTER 3

CHAPTER 6

List of Acronyms and Abbreviations

p16Ink4a - cyclin-dependent kinase inhibitor 2A
AAP: American Association of Pediatrics
AGP: α-1-acid glycoprotein
CALIPER: Canadian Laboratory Initiative on Pediatric Reference Intervals
CDC: Centers for Disease Control
CLSI: Clinical Laboratory Standards Institute
CIs: Confidence Intervals
CPS: Canadian Pediatric Society
CRP: C-reactive protein
ESPGHAN: European Society for Pediatric Gastroenterology, Hepatology and Nutrition
Figure Sxx: Supplemental figure number xx
GAMLSS: Generalized Additive Models for Location, Scale, and Shape
HiCN: Hemiglobincyanide
ID: Iron deficiency
IDA: Iron deficiency anemia
KiGGS: The German Health Interview and Examination Survey for Children and Adolescents
LOOK: Lifestyle Of Our Kids!
LMS: lambda-mu-sigma method
MCV: Mean Corpuscular Volume
NHANES: National Health and Nutrition Examination Survey
NAID: Non-anemic iron deficiency
NORICHILD: Scandinavian Initiative for the Establishment of Pediatric Reference Intervals
OCs: Optimal curves
ORACLE-FER: Optimized Reference Assessment for Clinical Laboratory Evaluation of Ferritin
ORACLE-H: Optimized Reference Assessment for Clinical Laboratory Evaluation of Hemoglobin
PRINCES-FERRITIN: Pediatric Reference Intervals and Curves Evidence Synthesis for Ferritin
PRINCES-H: Pediatric Reference Intervals and Curves Evidence Synthesis for Hemoglobin
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies
RCs: Reference curves
RIs: Reference intervals
RoB: Risk of bias
SES: Socio-economic status
SRM: Standardized reference material
Table Sxx: Supplemental table number xx
TARGet Kids!: The Applied Research Group for Kids!
WHO: World Health Organization

CHAPTER 1: INTRODUCTION AND OBJECTIVES

1.1. Reference intervals

Variation among humans in physiological processes due to, for example, differences in genetics, diseases, and environmental influences, are reflected in laboratory measurements. Accurate interpretations of laboratory results, therefore, require knowledge of the expected variation for a given laboratory measurement or analyte. The concept of defining expected variation dates back to the 19th century when in 1835, Adolphe Quetelet introduced the concept of the "average man".[1] More than one hundred years later, in 1969 a pivotal shift occurred when Gräsbeck and Saris challenged the concept of a universal "normal range" and introduced the modern framework of "reference intervals" (RIs), emphasizing interpretations relative to well-characterized reference populations rather than a fixed definition of normality.[1] RIs are traditionally established based on data from samples of individuals taken from healthy populations. Values between the lower and upper reference limits indicate what is considered "normal". The lower and upper reference limits typically correspond to the 2.5^{th} and 97.5^{th} percentiles of the reference distribution, respectively, hence the RI encompasses the middle 95% of the reference population.[2,3] Clinically, values outside these reference limits may signal a need for further evaluation.[4]

RIs are used by physicians and laboratory professionals for interpreting laboratory test results.[5,6] Over 70% of medical decisions are informed by results from laboratory tests.[7] The International Federation of Clinical Chemistry and Laboratory Medicine has published comprehensive sets of recommendations for the establishment of RIs, defining concepts such as reference limits and RIs, discussing the operational definition of "health," and providing detailed guidance on selecting reference populations, applying exclusion criteria, and partitioning by

factors such as age and sex. These recommendations are also outlined by the Clinical and Laboratory Standards Institute (CLSI) guidelines.[6] In addition, the CLSI guidelines address methodological considerations, including statistical approaches for calculating reference limits and the challenges of transferring RIs between laboratories. While the CLSI guidelines generally recommend that each laboratory establish its own RIs to account for the context of the local population, full adherence to these protocols has been described to be difficult due to the resources required[8]

In addition to RIs, clinical decision limits (CDLs) represent thresholds associated with disease status or adverse clinical outcomes. CDLs can be established based on percentiles of a reference population but also frequently involve analysis of associations with health outcomes and consideration of expert opinion. In clinical settings, a result outside the RI limits does not necessarily require intervention as it may still reflect physiological variation observed in a healthy reference population, i.e., a value that is in the top or bottom 2.5% of what is expected in the reference population. However, values outside a CDL may indicate a need for a decision regarding a clinical action.

1.2 Pediatric reference intervals

RIs are conventionally estimated by dividing the study population into discrete age and sex groups (which are often referred to as "partitions"). RIs are then calculated separately for each distinct group. Children are continuously growing from birth through adolescence. Consequently, more age partitions than for adult populations are often required to properly reflect the changes that occur with various biological parameters.[6,9,10] While some pediatric RIs are estimated by 5-year age intervals[11], others are estimated by 2 or 3-years age intervals.[12–14] Some pediatric RIs require partitioning in even smaller age groups, including less than 1-year

age intervals,[9,15,16] which is especially common for the first year of life,[9,17] when monthly or weekly partitioning may be required in order to provide accurate and reliable RIs. The large number of age partitions required in pediatric populations can lead to very sparse data, especially for younger children, leading to inadequate sample sizes.[18] Blood procurement from small children can also be difficult, because 10 mL of blood could constitute 10% of blood volume in this population and the added layer of obtaining parental consent to draw blood from their healthy child.[19]

1.3 Reference Curves

Separate RIs are often required for different sub-groups of the population, which can be achieved through partitioning, as described above.[6] This categorical partitioning may not fully reflect the continuous nature of biological variation during development. In pediatric populations, where biomarker concentrations often change rapidly with development, reliance on discrete age partitions can thus create artificial boundaries. As a result, children at the edges of an age group may be misclassified simply because of how partitions are assembled. Reference curves (RCs) consist of a series of smooth curves which represent centiles of the distributions of the measurements with respect to age.[20,21] With RCs, age is modeled as a continuous variable, which means RI estimates can be provided at any age. Therefore, RCs may better capture the true dynamics of the pediatric population during development.[20–22] Moreover, results for individuals do not suddenly change when they move from one age group to another. Hence, results for those who are close to the boundary of an age partition may be interpreted more accurately with RCs.[23] In addition, avoiding the need for extensive age partitioning addresses the limitation related to inadequate sample size in the pediatric population. Partitioning with respect to other variables

(e.g. sex, ethnicity, etc.) may still be needed depending on the nature of the biomarker, necessitating the estimation of separate RCs in some cases.

1.4 Statistical methods for estimation of RIs and RCs

1.4.1 Estimation of RIs

Reference intervals are typically established using samples from a healthy reference population, an approach known as the direct method of estimating reference intervals.[6] There are two general methods of selecting reference individuals from the reference population. When the criteria for selecting reference individuals are applied before samples are collected and analyzed, it is referred to as *a priori*. If the same criteria are applied following sample selection, it is referred as *a posteriori*. In addition, studies using indirect methods estimate reference intervals from large databases, such as hospital information systems or community based datasets. The database contains a mixture of measurements from both healthy and unhealthy individuals. Statistical techniques are then used to estimate the distribution of values representing the underlying healthy population within the overall dataset. These estimates are then used to derive reference intervals. The CLSI guidelines acknowledge the use of indirect methods. However, they also note reservations regarding these approaches and strongly encourage use of direct method.[6]

Once data from a healthy reference population is made available, there are four key steps involved in the establishment of RIs. These steps consist of identifying outliers, partitioning into homogenous groups, estimating the reference limits, and calculating the confidence intervals (CIs) (typically 90% CIs are reported [6]) corresponding to estimates of the reference limits.

Identifying and removing outliers is an important first step as some apparently healthy individuals may have an underlying disease that has not been detected or diagnosed. There are several methods available to detect outliers, of which two of the most commonly used are the

Dixon method and the Tukey method.[24–26] Both of these methods have been recommended in the CLSI guidelines.[6] The Dixon method compares the distance between a suspected outlier and its neighbour to a proportion of the distance between the suspected outlier and the opposite endpoint of the data. It can only be applied to one suspected outlier at a time. The Tukey method identifies outliers by evaluating both tails of the distribution simultaneously.[26] Observations that fall below the lower bound or above the upper bound are considered outliers and may be removed. These limits are defined as: min = Q1 - 1.5 × IQR and max = Q3 + 1.5 × IQR, where Q1 represents the first quartile, Q3 the third quartile, and IQR the interquartile range (IQR = Q3 - Q1). On the other hand, the adjusted Tukey method modifies these fences to allow for asymmetry in the distribution, creating non-symmetric boundaries that account for data skewness.[27] In addition, Dunn–Smyth method applies a model-based approach that uses randomized quantile residuals from a fitted model. Observations with residuals far from the expected distribution range are flagged as outliers.[28]

After screening for and removing outliers, data are partitioned into homogenous groups to reflect the changes that occur with various biological parameters. Although outlier detection is often performed at the initial stage, it is advisable to check again for outliers after partitioning is done. There are several ways to compute partitions. A common way is through defining intervals. For example, RIs can be calculated for every one-year age interval. Another way of creating partitions is through previous clinical knowledge. For example, age can be divided into the intervals that reflect well-known developmental stages, if the distribution of analyte concentration is likely to be associated with those stages. Partitions can also be determined through visual data inspection. For example, plotting the mean and standard deviation across age and gender can reveal distinct patterns that can guide the selection of age partitions. After

creating initial partitions, each partition should be tested against subsequent partitions to confirm whether these partitions should remain separate or be combined. The CLSI guideline suggests several approaches proposed by Harris et al., Lahti et al., and Sinton et al.[29–31] for evaluation of partitions. For example, the Harris et al.[26] method accomplishes this based on differences in means and variance between successive partitions.

Once data is checked for outliers and partitioning is appropriately done with respect to relevant factors, RIs are estimated for each partition, and CIs for the estimated limits of the RI are provided as a measure of their precision. There are three main approaches used to calculate RIs: the parametric, non-parametric, and robust methods. The non-parametric method is most commonly used; it is a simple way to calculate RIs empirically using ranks that does not require any assumptions regarding the distribution of the data.[6] The robust method uses an iterative process to compute a measure of the centre of a dataset.[6] The parametric method requires the data to have an underlying Gaussian distribution. Non-Gaussian data may be transformed to an approximate Gaussian distribution to permit the application of the parametric method.

The last key stage is estimation of CIs. CIs for the lower and upper limits of the RI can be estimated for the parametric method using formulas provided by Daily 2014 and Soldberg 1986[32,33], for non-parametric method using the order statistic (ranked observations)[32,6], and for the robust method using bootstrap percentiles.[34]

1.4.2 Estimation of RCs

While the CLSI guidelines do not specify procedures for estimating RCs, common practice follows some of the recommendations provided for RI estimation. This includes identifying outliers, often using Tukey et al.[23] based method; estimating the RCs; and generating 90% confidence intervals for the RCs, typically through bootstrap resampling. Because age is

modeled as a continuous variable, a separate step to partition data into homogeneous age groups is not required. Instead, RIs may be obtained from RCs at any age by deriving the corresponding lower and upper RIs limits from the biomarker age specific regression model.

Several statistical methods for the estimation of RCs exist, each with advantages and limitations. The methods range from non-parametric and semi-parametric to parametric. Quantile regression is a commonly used non-parametric method, where quantiles (e.g., centiles) of a specific continuous outcome are estimated as smoothed functions of age without the need to make any assumptions about the corresponding distributions.[35,36] Directly modeling the observed data without assuming any distribution makes this method more resistant to outliers compared to parametric methods.[37] However, a major concern associated with quantile regression is the possibility that centiles may cross as they are modeled independently of each other.[38,37] Schnabel & Eilers (2013)[39] proposed the quantile sheets method to minimize the centiles crossing, where curve centiles are estimated by simultaneously smoothing in direction of both the dependent variable (outcome measure) and independent variable (e.g., age).[39] Muggeo and colleagues (2013)[37] developed methods of non-crossing quantile regression that can be used to completely avoid centiles crossing.[37] Quantile regression can be implemented in R with several packages including *quantreg*[40] and *quantregGrowth*.[41]

One of the most common parametric methods for estimating RCs is fractional polynomial regression.[36,42] It provides estimates of the mean and standard deviation centile curves for a continuous outcome of interest as separate polynomial functions of an explanatory variable such as age. It assumes that the mean and standard deviation of the response variable vary smoothly as a function of age, and that at each age, the response variable follows the normal distribution.[36,42] The fractional polynomial method is implemented in R through *mfp* package.[43]

The Lambda (Box-Cox skewness transformation), Median (central tendency), Sigma (coefficient of variation) (LMS) method developed by Cole (1988)[44] and Cole & Green (1992)[45] is a commonly used semi-parametric method for RC estimation.[36,46] The LMS method is Cole's generalisation of van't Hof, Wit, and Roede's (1985) method[47], which suggested removing skewness at each age by power transformation of the response variable. The LMS method allows the three parameters (L, M, and S) to vary as a function of time (typically age as the time variable): L(t), M(t), and S(t).[36] General Additive Models for Location, Scale, and Shape (GAMLSS) are extensions of the LMS method with additional parameters that can model kurtosis.[48] GAMLSS methods are implemented in the R software *gamlss* package.[46]

1.5 Providing reliable pediatric RIs for laboratory markers in Canada

Proper analysis of data from healthy pediatric populations is vital to provide reference values for accurate and reliable interpretations of test results and their implications for a child's health. Nevertheless, providing accurate and reliable RIs for the pediatric population in Canada has been challenging. This was highlighted by the results of a national survey of Canadian clinical laboratories, which documented a lack of harmonization in pediatric RIs in laboratories across Canada.[49] The study indicated that many of the laboratories either used pediatric RIs from out-of-date textbooks or from unknown sources, or they relied on RIs established based on non-representative populations including RIs estimated based on data from adult populations.[49,50]

Pre-analytical and analytical aspects of analytes and their measurements represent another important determinant in the estimation of pediatric reference intervals.[6] These include biological factors as well as methodological factors such as sample handling and storage conditions, calibration procedures, analyzer type, and the reference materials used for assay standardization. Different laboratory analyzers and assay platforms may rely on distinct

analytical principles and calibration systems, which can introduce systematic differences in measured biomarker concentrations even when testing the same specimen. International standardized reference material was adopted to improve comparability of results across assays. This standardization enabled traceability of measurements to a recognized international reference standard. Despite ongoing efforts toward laboratory standardization, measurable between-method differences across assays and analyzer platforms persist for the same laboratory tests.[49]

1.6 Overview of iron deficiency and anemia

Hematology and biochemistry markers are the most commonly ordered laboratory tests in pediatrics and are often collected during routine examinations.[51,52] This thesis will, therefore, focus on these markers as we try to understand the broader issues around pediatric RIs. To limit the scope of the thesis and to have a clinically focused objective, we will specifically consider markers related to iron deficiency (ID) and iron deficiency anemia (IDA).

Iron is the fourth most common element on earth and is found in nature and in foods. It is involved in many physiological functions in the body. Inadequate iron levels can impact a number of essential body processes that require iron for their functioning, including hemoglobin synthesis[53], central nervous system development[52], electron transport and DNA synthesis[53,54], and protection from infection[52]. Iron levels can range from iron overload at one side of the spectrum, iron sufficiency representing the status of sufficient amounts of iron for maintaining normal physiological functions, ID without anemia (NAID) representing the status where iron levels are insufficient to maintain normal physiological functioning, to IDA at the other end of the spectrum representing inadequate numbers of red blood cells due to ID.[52,53] Studies have shown an association between NAID in infancy and neuro-cognitive impairments, with evidence of these impairments persisting into adulthood and being associated with poor emotional health

and failure to complete secondary school.[55,56] NAID represents an early stage of ID.[53] It has been shown that without treatment, NAID may progress to IDA.[52,53,57,58] Studies have shown that IDA may in turn cause weakness, headache as well as motor and cognitive impairment, and can lead to death.[59,60] ID and IDA affect over 750 million children worldwide.[52,53,61] In the United States 3% of children below 36 months of age are affected by IDA compared with less then 1% of children between 37 to 60 months of age.[62] In Canada, the prevalence of IDA for children in general is between 3.5% and 10.5%[58], while some population groups exhibit higher prevalence. For example, in some Northern Ontario First Nations communities and some Inuit communities, the prevalence of IDA has been estimated at 36% and 56%, respectively.[61,63]

1.7 Overview of biomarkers – hemoglobin and serum ferritin

Hemoglobin (a hematology marker) and serum ferritin (a biochemistry marker) are used in screening for iron deficiency in children.[52] The assessment of these biomarkers is challenging due to considerable variations in the biomarkers' concentration levels particularly during early development, which can complicate the interpretation of laboratory results.[52] This often leads to gaps in clinical understanding that increases the risk of delayed or inaccurate diagnoses.[64] We, therefore, focus on these markers as we try to understand the broader issues around pediatric RIs.

1.7.1 Hemoglobin and its RIs

Anemia is generally defined as a condition characterized by an insufficient number of healthy red blood cells or an inadequate concentration of healthy hemoglobin to effectively deliver oxygen to body tissues.[65] Anemia can arise from causes including iron deficiency, hemolytic anemias, anemia of chronic disease, heavy menstrual bleeding, and anemia related nutrient deficiencies other than iron.[52]

Hemoglobin is an intra-cellular protein found in red blood cells whose main purpose is oxygen transport and indirect transport of carbon dioxide. It is a tetrameric protein composed of two α-globin and two β-globin subunits, each carrying a heme prosthetic group with a central iron atom. The structural transitions between the tense and relaxed state facilitate oxygen loading in the lungs and unloading in peripheral tissues.[66] Mutations or deletions in α-globin and β-globin genes underlie hemoglobinopathies such as sickle cell disease and thalassemias.[67] Historically, diagnosing anemia was largely clinical via symptoms including pallor, changes in the nail beds (such as spooning), and glossitis (red tongue).[68] Today, a hemoglobin test, typically part of a complete blood count, measures the concentration of hemoglobin in the blood.

Pediatric RIs for hemoglobin are essential tools for determining whether a child's hemoglobin concentration falls within the expected range for their age and sex. These intervals help guide interpretation of lab results and inform clinical decisions around follow-up, diagnosis, and treatment. However, hemoglobin RIs vary significantly across regions due to geographic, demographic, nutritional, and methodological differences.[69] For example, in Mozambican toddlers aged 1–2 years, lower reference limits have been reported as low as 68 g/L, with similarly low values observed in Gambian children of the same age (~68 g/L).[69] In contrast, studies conducted in North America, Europe, and Canada have reported higher RIs for children aged 1–5 years, with lower bounds closer to 107–109 g/L.[69] The lower reference limits for hemoglobin RIs have been documented to vary both below and above the widely used WHO thresholds for anemia.[69] Importantly, this discrepancy is partly methodological: WHO defines anemia thresholds using the 5th percentile of the distribution in healthy populations, whereas RIs typically capture the central 95% of this distribution, from the 2.5th to 97.5th percentiles. As a

result, the lower limit of an RI is expected to fall below the WHO threshold simply due to how the boundaries are defined.[6]

1.7.2 Ferritin and its RIs

Ferritin is an intracellular iron storage protein. Since ferritin also circulates in serum and plasma, it is classified as a biochemical biomarker, in contrast to hemoglobin, which is considered a hematological biomarker of red cells. Ferritin concentration in blood serum and plasma serves as a key biomarker for assessing iron status.[70,71] The gold standard for diagnosing ID status is the assessment of bone marrow iron content through iron staining while the assessment of liver iron content through biopsy and measurement of iron concentration is important for diagnosing iron overload.[72] These procedures are highly invasive and costly, which limits their use in clinical settings. As an alternative, measurements of ferritin concentrations are used as indirect measures for the assessment of iron status.[72] Pediatricians can compare laboratory measurements of serum ferritin concentrations to existing ferritin RIs that describe the distribution of ferritin concentration in a healthy reference population.[73,74] However, there are inconsistencies across studies in their estimates of RIs for serum ferritin among children of the same age.[74] For example, in the Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER) study based in Canada, the lower reference limit for ferritin in children aged 1 to <5 years was reported for both sexes as 5.3 µg/L, with an upper limit around 99.9 µg/L.[75] Meanwhile, RI estimates obtained based on data from German populations showed for children aged 3 years, the lower limit ranged from 11.9 to 13.8 µg/L for boys and girls, with upper limits ranging from 88 to 87.2 µg/L.[76] Depending on the study, a ferritin value of 10 µg/L in a toddler might be considered within the RI based on Canadian estimates but below normal with respect to estimates from the German data. Additionally, the WHO defines iron deficiency in children aged

6–59 months as ferritin < 12 µg/L.[71] In contrast, the 2025 American Society of Hematology (ASH) draft recommendations suggest a ferritin threshold of ≤20 µg/L in children 9 months to 4 years.[77] These inconsistencies complicate the interpretation of ferritin laboratory results and create uncertainty for clinicians and in turn for children and their families.

1.8 Optimal curves for hemoglobin and serum ferritin

1.8.1 Concept of optimal curves

Traditionally RIs are developed using a sample from healthy individuals, referred to as the reference population. An alternative to the conventional approach using a broad sample of apparently healthy individuals is to define the reference population focusing on those with characteristics favourable to achieving optimal biomarker status. This concept, described in anthropometric research as the development of "reference standards", involves selecting a sub-sample of individuals free from health conditions, and without environmental or socioeconomic constraints that could impair the biomarker of interest.[78,79] As such, the resulting curves represent not merely the range of values observed in the general healthy population but rather a prescriptive reference standard, a target biomarker levels that reflects the physiological potential under optimal conditions. While this methodology is well established in anthropometric measurements for children such as height and weight, it has not been systematically applied in laboratory medicine.

1.8.2 Iron optimal intervals and curves

The creation of such optimal curves for biochemical or hematologic parameters would provide a benchmark for both individual clinical assessment and population level evaluation. In the context of iron, several biological, environmental, and social factors are known to influence iron status; these determinants are reviewed next to inform the selection of an optimal reference population.

Prematurity and low birth weight are known risk factors for reduced iron stores at birth. The Canadian Paediatric Society defines prematurity as birth before 37 weeks' gestation and low birth weight as less than 2,500 g.[53] These factors are associated with incomplete iron transfer from mother to child, which occurs primarily during the third trimester.[53] As a result, infants born preterm and/or with low birth weight are at increased risk of ID.[53] In contrast, full-term infants are generally born with sufficient iron stores to meet their needs for the first four to six months of life.[52,53]

Socio-economic status has been consistently shown to be associated with iron status in young children across multiple studies. Thane et al., (2000)[80], analyzing data from British toddlers aged 1.5 to 4.5 years, found that children from lower income households were at significantly higher risk for poor iron status, as indicated by lower hemoglobin and serum ferritin concentrations. The study suggested that dietary factors, such as lower intake of iron rich or fortified foods, may mediate this relationship. Similarly, Baker and Greer (2010)[52] highlighted that children from economically disadvantaged backgrounds are at increased risk for ID, largely due to limited access to iron-rich foods, inadequate nutrition education, and higher rates of food insecurity. Bayoumi et al., (2020)[81] further strengthened this evidence by analyzing data from the TARGet Kids! cohort and showing an association between family income, food insecurity, and iron status among young Canadian children. Children from low income families had higher odds of low serum ferritin, even after adjusting for dietary iron intake and other covariates.[81]

Inflammatory markers such as C-reactive protein (CRP) and α-1-acid glycoprotein (AGP) are important for interpreting ferritin levels, which can become elevated in the presence of infection or inflammation, potentially masking underlying ID.[82,83] Both the American Association of Pediatrics and the WHO recommend measuring CRP and/or AGP alongside

ferritin to account for this effect.[52,71] CRP is more commonly used due to its greater availability in clinical settings.[84,85]

Multiple studies report that underweight, overweight and obese children are at increased risk for ID. A Canadian study led by Borkhoff et al. (2023)[86] demonstrated that ID was more prevalent not only among underweight children but also among those who are overweight or obese. In a Brazilian study, Bagni et al. (2012)[87] found that adolescent girls who are overweight or obese had significantly lower hemoglobin levels compared to non overweight girls. Similar findings were reported in studies from Saudi Arabia and Egypt, which showed a higher prevalence of ID and IDA among obese children.[88,89]

Another risk factor identified by the Canadian Paediatric Society (CPS) is exclusive breastfeeding beyond six months of age.[51] Infant iron stores typically meet physiological requirements only during the first six months of life, and breast milk contains relatively low concentrations of iron.[51,55] Consequently, prolonged exclusive breastfeeding without appropriate iron supplementation may increase the risk of iron deficiency. In addition, studies have shown that maternal anemia, hypertension, and diabetes may also influence neonatal iron status.[50]

1.9 Available databases and cohorts

Development of RIs, RCs, OCs and corresponding optimal intervals (OIs, which can be derived from OCs) require the selection of a cohort of individuals who are representative of the population and provide measurements of the biomarker of interest across the relevant age span, including data on other characteristics needed for partitioning and for selection of subgroups with characteristics associated with optimal iron status for OC estimation.

Sources of data on healthy children for the purpose of establishing RIs and RCs exist both globally and within Canada. For instance, the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) provides nationwide representative cross-sectional data for the German pediatric population.[90] Similarly, the Australian Harmonising Age Pathology Parameters in Kids! (HAPPI Kids!) study collects cross-sectional blood sample data from children aged 30 days to 18 years with aim of developing age specific RIs for the most commonly tested biomarkers.[91] In Scandinavia, the NORICHILD initiative, a collaborative effort between Finland, Sweden, and Denmark, aims to establish harmonized pediatric RIs for 25 extensively used biomarkers.[92,93] There are several Canadian cohorts and initiatives dedicated to studying to the study of healthy pediatric populations. CALIPER, for instance, is a major initiative established to provide pediatric RIs for wide range of biomarkers. It is a population-based study that collects pediatric data directly from the community. On the other hand, the Canadian Health Measures Survey (CHMS) is a national, cross-sectional study that collects health related data on Canadians aged 3 to 79 years.[94,95] The Applied Research Group for Kids! (TARGet Kids!) study collects both cross-sectional and longitudinal health data from a diverse cohort of healthy Canadian children from age 2 weeks and older, including data on children collected both from children and their parents.[96] The TARGet Kids! cohort also reflects an ethnically and socioeconomically diverse Canadian pediatric population.[96]

As new data become available from ongoing studies that have been used to estimate RIs and RCs, there is an opportunity to update the estimates to improve accuracy and precision by making best use of data as soon as it becomes available. This strategy is conceptually similar to the idea of a living systematic review and meta-analysis, an emerging approach that involves updating systematic reviews and the pooled estimates as new primary studies become

available.[97,98] Similarly, the RI estimates obtained from the analysis of data from a healthy cohort of children can also be updated as more data become available. As new knowledge for maintaining optimal iron status becomes available, this information could be incorporated into the "living" OCs and OIs. Updating the estimates with the most up-to-date data and emerging evidence also aligns with WHO recommendations that RIs should be regularly updated.[99]

1.10 Thesis objectives

This thesis aims to evaluate current practices in establishing pediatric RIs and RCs, develop RIs, RCs, OIs and OCs and provide methodological and computational frameworks for synthesizing evidence and generating more accurate and reliable reference and optimal intervals. The specific objectives are:

1) to synthesize evidence on pediatric RIs and RCs for hemoglobin and ferritin and examine heterogeneity to identify its potential sources (Chapters 2-4).

2) to establish pediatric RCs for hemoglobin and ferritin using data from Canadian children and generate age- and sex-specific RIs based on the estimated RCs (Chapters 5-6).

3) to establish pediatric OCs for hemoglobin and ferritin using the same data source as that used for objective (2) and generate age- and sex-specific OIs based on OCs (Chapters 5-6).

4) to create interactive web-based computational and graphical tools related to objectives 1 through 3 (Chapters 3-6).

The method used to pursue the objective to synthesize published RIs/RCs from the current literature (objective 1) is a systematic review and meta-analysis. The analysis of cross-sectional data from a Canadian cohort of children was used to pursue objectives two and three. To enhance

the accessibility and applicability of our findings, these objectives are accompanied with an interactive Shiny web application (objective 4) allowing researchers, clinicians, and health policy makers to explore and apply the results dynamically.

1.11 Thesis organisation

This manuscript-based thesis is organized in accordance with the School of Epidemiology and Public Health (SEPH), University of Ottawa, Guidance for Thesis by Article. The following five consecutive chapters (Chapters 2-6) have each been formatted as manuscripts for publication, with Chapter 7 presenting an integrated discussion.

Specifically, Chapters 2-4 present manuscripts addressing the first objective, including the combined study protocol for hemoglobin and ferritin systematic reviews, followed by the results from each systematic review and meta-analysis:

- Chapter 2 presents manuscript 1, titled "Pediatric reference intervals and curves for hemoglobin and ferritin: protocol for a systematic review and meta-analysis" published in *BMJ Open.*
- Chapter 3 presents manuscript 2, titled "Pediatric Reference Intervals and Curves for Hemoglobin Estimated Using Direct Methods: A Systematic Review and Meta-Analysis" published in *International Journal of Laboratory Hematology*.
- Chapter 4 presents manuscript 3, titled "Pediatric Reference Intervals and Curves for Ferritin Estimated Using Direct Methods: A Systematic Review and Meta-Analysis" under revision for publication to *BMC Pediatrics*.

Chapters 5-6 present manuscripts addressing the second and third thesis objectives, including the estimated pediatric RIs, RCs, OIs, and OCs for pediatric hemoglobin and ferritin:

- Chapter 5 presents manuscript 4, titled “Beyond reference intervals: optimizing lower and upper limits for reporting normative hemoglobin levels for children” under revision for publication to *JAMA Network Open* journal.
- Chapter 6 presents paper 5, titled “Ferritin reference and optimal curves compared with thresholds in pre-adolescents” accepted with minor revisions for publication to *JAMA Network Open* journal.

Finally, Chapter 7 draws upon the findings from the earlier six chapters and presents an integrated discussion of the key findings and conclusions from all of the component articles and the thesis as a whole, including recommendations to guide future work.

References

1 El-Khoury JM, Badrick T, Theodorsson E. Time to Reevaluate the 95% Inclusion Criteria for Defining Reference Intervals? *Clin Chem* 2024; **70**: 700–2.

2 Aldrimer M, Ridefelt P, Rödöö P, Niklasson F, Gustafsson J, Hellberg D. Population-based pediatric reference intervals for hematology, iron and transferrin. *Scand J Clin Lab Invest* 2013; **73**: 253–61.

3 Shiefa S, Amargandhi M, Bhupendra J, Moulali S, Kristine T. First Trimester Maternal Serum Screening Using Biochemical Markers PAPP-A and Free β-hCG for Down Syndrome, Patau Syndrome and Edward Syndrome. *Indian J Clin Biochem* 2013; **28**: 3–12.

4 Daly CH, Higgins V, Adeli K, Grey VL, Hamid JS. Reference interval estimation: Methodological comparison using extensive simulations and empirical data. *Clin Biochem* 2017; **50**: 1145–58.

5 Ceriotti F, Hinzmann R, Panteghini M. Reference intervals: the way forward. *Ann Clin Biochem Int J Lab Med* 2009; **46**: 8–17.

6 Clinical and Laboratory Standards Institute (CLSI), editor. Defining, Establishing and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition. Wayne, Pa: Clinical and Laboratory Standards Institute, 2008.

7 Ngo A, Gandhi P, Miller WG. Frequency that Laboratory Tests Influence Medical Decisions. *J Appl Lab Med* 2017; **1**: 410–4.

8 Ozarda Y. Reference intervals: current status, recent developments and future considerations. *Biochem Medica* 2016; : 5–16.

9 Daly CH, Liu X, Grey VL, Hamid JS. A systematic review of statistical methods used in constructing pediatric reference intervals. *Clin Biochem* 2013; **46**: 1220–7.

10 Adeli K, Higgins V, Trajcevski K, White-Al Habeeb N. The Canadian laboratory initiative on pediatric reference intervals: A CALIPER white paper. *Crit Rev Clin Lab Sci* 2017; **54**: 358–413.

11 Garel C, Cont I, Alberti C, Josserand E, Moutard ML, Ducou Le Pointe H. Biometry of the Corpus Callosum in Children: MR Imaging Reference Data. *Am J Neuroradiol* 2011; **32**: 1436–43.

12 Clifford SM, Bunker AM, Jacobsen JR, Roberts WL. Age and gender specific pediatric reference intervals for aldolase, amylase, ceruloplasmin, creatine kinase, pancreatic amylase, prealbumin, and uric acid. *Clin Chim Acta* 2011; **412**: 788–90.

13 Xu S, Morse AM, Lacy B, Baggett LS, Gogola GR. Peg Restrained Intrinsic Muscle Evaluator (PRIME): Development, Reliability, and Normative Values of a Device to Quantify Intrinsic Hand Muscle Strength in Children. *J Hand Surg* 2011; **36**: 894–903.

14 Tamimi W, Tamim H, Felimban N, AlMutair A, Altwaijri Y, AlAlwan I. Age- and gender-specific reference intervals for serum lipid levels (measured with an Advia 1650 analyzer) in school children. *Pediatr Int* 2011; **53**: 814–9.

15 Uemura O, Honda M, Matsuyama T, *et al.* Age, gender, and body length effects on reference serum creatinine levels determined by an enzymatic method in Japanese children: a multicenter study. *Clin Exp Nephrol* 2011; **15**: 694–9.

16 Hamid JS, Atenafu EG, Borkhoff CM, *et al.* Reference intervals for hemoglobin and mean corpuscular volume in an ethnically diverse community sample of Canadian children 2 to 36 months. *BMC Pediatr* 2021; **21**: 241.

17 Wang B, Zhang J, Wu Y-Z, Lu Z-H, Wang N, Yu Z-B. Reference Interval for Pulse Oxygen Saturation in Neonates at Different Altitudes: A Systematic Review. *Front Pediatr* 2021; **9**: 771750.

18 Zierk J, Baum H, Bertram A, *et al.* High-resolution pediatric reference intervals for 15 biochemical analytes described using fractional polynomials. *Clin Chem Lab Med CCLM* 2021; **59**: 1267–78.

19 Clejan S. Neonatology & Laboratory Medicine. Anne Green, Imogen Morgan, and Jim Gray. London, UK: ACB Venture Publications, 2003, 328 pp., £24. ISBN 0-902429-41-8. *Clin Chem* 2004; **50**: 460–460.

20 Rosner B, Cook N, Portman R, Daniels S, Falkner B. Determination of Blood Pressure Percentiles in Normal-Weight Children: Some Methodological Issues. *Am J Epidemiol* 2008; **167**: 653–66.

21 Brouwer MJ, De Vries LS, Groenendaal F, *et al.* New Reference Values for the Neonatal Cerebral Ventricles. *Radiology* 2012; **262**: 224–33.

22 Faulkner WR, Demers LM. Importance of age-dependent reference values in biochemical testing: are we including the elderly population? *Clin Chem* 1994; **40**: 855–6.

23 Rahman M, George C, Monagle P. Hot topics in coagulation testing: Important considerations for testing children for bleeding/thrombotic disorders. *Int J Lab Hematol* 2020; **42**: 68–74.

24 Dixon WJ. Processing Data for Outliers. *Biometrics* 1953; **9**: 74.

25 Reed AH, Henry RJ, Mason WB. Influence of statistical method used on the resulting estimate of normal range. *Clin Chem* 1971; **17**: 275–84.

26 Tukey JW. Exploratory data analysis. Reading, Mass: Addison-Wesley Pub. Co, 1977.

27 Hubert M, Van Der Veeken S. Outlier detection for skewed data. *J Chemom* 2008; **22**: 235–46.

28 Dunn PK, Smyth GK. Randomized Quantile Residuals. *J Comput Graph Stat* 1996; **5**: 236–44.

29 Harris EK, Boyd JC. On dividing reference data into subgroups to produce separate reference ranges. *Clin Chem* 1990; **36**: 265–70.

30 Lahti A. Partitioning biochemical reference data intosubgroups: comparison of existing methods. *Clin Chem Lab Med CCLM* 2004; **42**. DOI:10.1515/CCLM.2004.123.

31 Sinton TJ, Cowley DM, Bryant SJ. Reference intervals for calcium, phosphate, and alkaline phosphatase as derived on the basis of multichannel-analyzer profiles. *Clin Chem* 1986; **32**: 76–9.

32 Daly CH. Methods for Estimating Reference Intervals. 2014. http://hdl.handle.net/11375/15398.

33 Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *Clin Chim Acta* 1987; **170**: S13–32.

34 Efron B, Tibshirani RJ. An Introduction to the Bootstrap, 0 edn. Chapman and Hall/CRC, 1994 DOI:10.1201/9780429246593.

35 Koenker R, Chernozhukov V, He X, Peng L, editors. Handbook of Quantile Regression, 1st edn. Chapman and Hall/CRC, 2017 DOI:10.1201/9781315120256.

36 Ohuma EO, Altman DG, for the International Fetal and Newborn Growth Consortium for the 21 Century (INTERGROWTH-21 Project). Statistical methodology for constructing gestational age-related charts using cross-sectional and longitudinal data: The INTERGROWTH-21st project as a case study. *Stat Med* 2019; **38**: 3507–26.

37 Muggeo VMR, Sciandra M, Tomasello A, Calvo S. Estimating growth charts via nonparametric quantile regression: a practical framework with application in ecology. *Environ Ecol Stat* 2013; **20**: 519–31.

38 Wei Y, Pere A, Koenker R, He X. Quantile regression methods for reference growth charts. *Stat Med* 2006; **25**: 1369–82.

39 Schnabel SK, Eilers PHC. Simultaneous estimation of quantile curves using quantile sheets. *AStA Adv Stat Anal* 2013; **97**: 77–87.

40 Koenker R. quantreg: Quantile Regression. 2024. https://CRAN.R-project.org/package=quantreg.

41 Muggeo VMR, Torretta F, Eilers PHC, Sciandra M, Attanasio M. Multiple smoothing parameters selection in additive regression quantiles. *Stat Model* 2021; **21**: 428–48.

42 Royston P, Altman DG. Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. *Appl Stat* 1994; **43**: 429.

43 Ambler G, Benner A. mfp: Multivariable Fractional Polynomials. 2024. https://CRAN.R-project.org/package=mfp.

44 Cole TJ. Fitting Smoothed Centile Curves to Reference Data. *J R Stat Soc Ser A Stat Soc* 1988; **151**: 385.

45 Cole TJ, Green PJ. Smoothing reference centile curves: The lms method and penalized likelihood. *Stat Med* 1992; **11**: 1305–19.

46 Stasinopoulos MD, Rigby RA, Heller GZ, Voudouris V, Bastiani FD. Flexible Regression and Smoothing: Using GAMLSS in R, 1st edn. Chapman and Hall/CRC, 2017 DOI:10.1201/b21973.

47 van 't Hof MA, Wit JM, Roede MJ. A method to construct age references for skewed skinfold data, using Box-Cox transformations to normality. *Hum Biol* 1985; **57**: 131–9.

48 Rigby RA, Stasinopoulos DM. Smooth centile curves for skew and kurtotic data modelled using the Box–Cox power exponential distribution. *Stat Med* 2004; **23**: 3053–76.

49 Adeli K, Higgins V, Seccombe D, *et al.* National Survey of Adult and Pediatric Reference Intervals in Clinical Laboratories across Canada: A Report of the CSCC Working Group on Reference Interval Harmonization. *Clin Biochem* 2017; **50**: 925–35.

50 Higgins V. What is the current state of reference intervals in Canada? A need for harmonization. 2017; published online Nov 20. https://myadlm.org/science-and-research/scientific-shorts/2017/current-state-of-reference-intervals-in-canada.

51 George-Gay B, Parker K. Understanding the complete blood count with differential. *J Perianesth Nurs* 2003; **18**: 96–117.

52 Baker RD, Greer FR, The Committee on Nutrition. Diagnosis and Prevention of Iron Deficiency and Iron-Deficiency Anemia in Infants and Young Children (0–3 Years of Age). *Pediatrics* 2010; **126**: 1040–50.

53 Unger SL, Fenton TR, Jetty R, Critch JN, O'connor DL. Iron requirements in the first 2 years of life. *Paediatr Child Health* 2019; **24**: 555–555.

54 Rao R, Georgieff MK. Iron in fetal and neonatal nutrition. *Semin Fetal Neonatal Med* 2007; **12**: 54–63.

55 Lozoff B, Smith JB, Kaciroti N, Clark KM, Guevara S, Jimenez E. Functional Significance of Early-Life Iron Deficiency: Outcomes at 25 Years. *J Pediatr* 2013; **163**: 1260–6.

56 Parkin PC, Borkhoff CM. Practical tips for paediatricians: Assessment and management of young children with iron deficiency. *Paediatr Child Health* 2018; **23**: 433–4.

57 Abdullah K, Thorpe KE, Maguire JL, *et al.* Risk factors, practice variation and hematological outcomes of children identified with non-anemic iron deficiency following screening in primary care setting. *Paediatr Child Health* 2015; **20**: 302–6.

58 Lozoff B, Jimenez E, Hagen J, Mollen E, Wolf AW. Poorer Behavioral and Developmental Outcome More Than 10 Years After Treatment for Iron Deficiency in Infancy. *Pediatrics* 2000; **105**: e51–e51.

59 Mireku MO, Davidson LL, Koura GK, *et al.* Prenatal Hemoglobin Levels and Early Cognitive and Motor Functions of One-Year-Old Children. *Pediatrics* 2015; **136**: e76–83.

60 Scott S, Chen-Edinboro L, Caulfield L, Murray-Kolb L. The Impact of Anemia on Child Mortality: An Updated Review. *Nutrients* 2014; **6**: 5915–32.

61 Christofides A, Schauer C, Zlotkin SH. Iron deficiency anemia among children: Addressing a global public health problem within a Canadian context. *Paediatr Child Health* 2005; **10**: 597–601

62 Centers for Disease Control (CDC). CDC criteria for anemia in children and childbearing-aged women. *MMWR Morb Mortal Wkly Rep* 1989; **38**: 400–4.

63 Hartfield D. Iron deficiency is a public health problem in Canadian infants and children. *Paediatr Child Health* 2010; **15**: 347–50.

64 Wynn R. Pediatric Hematology, 1st ed. West Nyack: Cambridge University Press, 2017.

65 Mayo Foundation for Medical Education and Research (MFMER). Anemia. 2026. https://www.mayoclinic.org/diseases-conditions/anemia/symptoms-causes/syc-20351360 (accessed March 15, 2026).

66 Monod J, Wyman J, Changeux J-P. On the nature of allosteric transitions: A plausible model. *J Mol Biol* 1965; **12**: 88–118.

67 Weatherall DJ. The inherited diseases of hemoglobin are an emerging global health burden. *Blood* 2010; **115**: 4331–6.

68 Walker HK, Hall WD, Hurst JW, editors. Chapter 151. In: Clinical methods: the history, physical, and laboratory examinations, 3rd ed. Boston: Butterworths, 1990.

69 Jorgensen JM, Crespo-Bellido M, Dewey KG. Variation in hemoglobin across the life cycle and between males and females. *Ann N Y Acad Sci* 2019; : nyas.14096.

70 Knovich MA, Storey JA, Coffman LG, Torti SV, Torti FM. Ferritin for the clinician. *Blood Rev* 2009; **23**: 95–104.

71 World Health Organization. WHO guideline on use of ferritin concentrations to assess iron status in individuals and populations. Geneva: World Health Organization, 2020 https://iris.who.int/handle/10665/331505 (accessed June 18, 2025).

72 Garcia-Casal MN, Pasricha S-R, Martinez RX, Lopez-Perez L, Peña-Rosas JP. Are Current Serum and Plasma Ferritin Cut-offs for Iron Deficiency and Overload Accurate and Reflecting Iron Status? A Systematic Review. *Arch Med Res* 2018; **49**: 405–17.

73 Ozarda Y, Sikaris K, Streichert T, Macri J, on behalf of IFCC Committee on Reference intervals and Decision Limits (C-RIDL). Distinguishing reference intervals and clinical decision limits – A review by the IFCC Committee on Reference Intervals and Decision Limits. *Crit Rev Clin Lab Sci* 2018; **55**: 420–31.

74 Eie AMØ, Strand MF, Fredriksen PM, Lindberg M. Population-based reference intervals for ferritin, iron, transferrin and transferrin saturation and prevalence of iron deficiency in 6–12-year-old children: the Health Oriented Pedagogical Project (HOPP). *Scand J Clin Lab Invest* 2021; **81**: 208–12.

75 Bailey D, Colantonio D, Kyriakopoulou L, *et al.* Marked Biological Variance in Endocrine and Biochemical Markers in Childhood: Establishment of Pediatric Reference Intervals Using Healthy Community Children from the CALIPER Cohort. *Clin Chem* 2013; **59**: 1393–405.

76 Rieger K, Vogel M, Engel C, *et al.* Reference intervals for iron-related blood parameters: results from a population-based cohort study (LIFE Child). *LaboratoriumsMedizin* 2016; **40**: 000010151520160019.

77 The American Society of Hematology (ASH). ASH Draft Recommendations for Diagnosis of Iron Deficiency. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwj-h7em0LiQAxWf1fACHbWnMtcQFnoECBwQAQ&url=https%3A%2F%2Fwww.hematology.org%2F-%2Fmedia%2Fhematology%2Ffiles%2Feducation%2Fclinicians%2Fguidelines-quality%2Fida-dx-public-comment.pdf.

78 De Onis M, Garza C, Habicht J-P. Time for a New Growth Reference. *Pediatrics* 1997; **100**: e8–e8.

79 de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. *Food Nutr Bull* 2004; **25**: S15-26.

80 Thane C, Walmsley C, Bates C, Prentice A, Cole T. Risk factors for poor iron status in British toddlers: further analysis of data from the National Diet and Nutrition Survey of children aged 1.5–4.5 years. *Public Health Nutr* 2000; **3**: 433–40.

81 Bayoumi I, Parkin PC, Birken CS, Maguire JL, Borkhoff CM, for the TARGet Kids! Collaboration. Association of Family Income and Risk of Food Insecurity With Iron Status in Young Children. *JAMA Netw Open* 2020; **3**: e208603.

82 Beard JL, Murray-Kolb LE, Rosales FJ, Solomons NW, Angelilli ML. Interpretation of serum ferritin concentrations as indicators of total-body iron stores in survey populations: the role of biomarkers for the acute phase response. *Am J Clin Nutr* 2006; **84**: 1498–505.

83 Knovich MA, Storey JA, Coffman LG, Torti SV, Torti FM. Ferritin for the clinician. *Blood Rev* 2009; **23**: 95–104.

84 Parkin PC, Hamid J, Borkhoff CM, *et al.* Laboratory reference intervals in the assessment of iron status in young children. *BMJ Paediatr Open* 2017; **1**: e000074.

85 Namaste SM, Rohner F, Huang J, *et al.* Adjusting ferritin concentrations for inflammation: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *Am J Clin Nutr* 2017; **106**: 359S-371S.

86 Borkhoff SA, Parkin PC, Birken CS, Maguire JL, Macarthur C, Borkhoff CM. Examining the Double Burden of Underweight, Overweight/Obesity and Iron Deficiency among Young Children in a Canadian Primary Care Setting. *Nutrients* 2023; **15**: 3635.

87 Bagni UV, Luiz RR, Veiga GVD. Overweight is associated with low hemoglobin levels in adolescent girls. *Obes Res Clin Pract* 2013; **7**: e218–29.

88 Saad RA, Qutob HM. The relationship between anemia and obesity. *Expert Rev Hematol* 2022; **15**: 911–26.

89 Sharourou ASA, Hassan MA, Teclebrhan MB. Anemia : Its Prevalence, Causes and Management. *Egypt J Hosp Med* 2018; **70**: 1877–9.

90 Hoffmann R, Lange M, Butschalowsky H, *et al.* KiGGS Wave 2 cross-sectional study – participant acquisition, response rates and representativeness. 2018; published online March 15. DOI:10.25646/5562.

91 Hoq M, Karlaftis V, Mathews S, *et al.* A prospective, cross-sectional study to establish age-specific reference intervals for neonates and children in the setting of clinical biochemistry, immunology and haematology: the HAPPI Kids study protocol. *BMJ Open* 2019; **9**: e025897.

92 Hoq M, Matthews S, Donath S, Carlin J, Ignjatovic V, Monagle P. Paediatric Reference Intervals: Current Status, Gaps, Challenges and Future Considerations. *Clin Biochem Rev* 2020; **41**: 43–52.

93 Rustad P, Felding P, Franzson L, *et al.* The Nordic Reference Interval Project 2000: recommended reference intervals for 25 common biochemical properties. *Scand J Clin Lab Invest* 2004; **64**: 271–84.

94 Cooper M, Greene-Finestone L, Lowell H, Levesque J, Robinson S. Iron sufficiency of Canadians. *Health Rep* 2012; **23**: 41–8.

95 Tremblay MS, Gorber SC. Canadian Health Measures Survey: Brief Overview. *Can J Public Health* 2007; **98**: 453–6.

96 Carsley S, Borkhoff CM, Maguire JL, *et al.* Cohort Profile: The Applied Research Group for Kids (TARGet Kids!). *Int J Epidemiol* 2015; **44**: 776–88.

97 BMJ. Living systematic reviews: towards real-time evidence for health-care decision making. 2025 https://bestpractice.bmj.com/info/toolkit/discuss-ebm/living-systematic-reviews-towards-real-time-evidence-for-health-care-decision-making/ (accessed June 18, 2025).

98 Cochrane Community. Living Systematic Review. 2025 https://community.cochrane.org/review-production/production-resources/living-systematic-reviews (accessed June 18, 2025).

99 Khadilkar V, Khadilkar A. Growth charts: A diagnostic tool. *Indian J Endocrinol Metab* 2011; **15**: 166.

CHAPTER 2: PEDIATRIC REFERENCE INTERVALS AND CURVES FOR HEMOGLOBIN AND FERRITIN: PROTOCOL FOR A SYSTEMATIC REVIEW AND META-ANALYSIS

2.1. Preface to Chapter 2

The first objective of this thesis is to addresses the limited understanding of the existing knowledge concerning pediatric reference intervals for commonly used hematological and biochemical biomarkers hemoglobin and serum and plasma ferritin. Specifically, our aim is to identify and synthesize available literature on these selected hematological and biochemical markers including describing what data sources and types are being used, quantify heterogeneity across various studies establishing reference intervals, and identify and elucidate sources of heterogeneity. To our knowledge no comprehensive systematic assessment of pediatric reference intervals for this group of analytes have been conducted to date.

This manuscript corresponding to this chapter (a protocol to two systematic reviews) is published in *BMJ Open*. We published the protocol to ensure transparency, reproducibility, and methodological rigor in the systematic review and meta-analysis research process. By making the protocol publicly available, we have enabled peer review scrutiny, provided a basis for methodological comparison for the final completed review, and helped prevent unnecessary duplication of effort by informing the peers of ongoing work.

VB was involved in the design and writing of the initial draft of the protocol. JSH, BP, PP, FM, and ML were involved in the design, writing, and revision of the protocol.

2.2 **Manuscript status:** published in *BMJ Open*

Paediatric reference intervals and curves for haemoglobin and ferritin: protocol for a systematic review and meta-analysis

Bijelić, V., Potter, B., Parkin, P. C., Momoli, F., Liebman, M., & Hamid, J. S. (2023). Paediatric reference intervals and curves for haemoglobin and ferritin: protocol for a systematic review and meta-analysis. *BMJ open*, *13*(10), e073783. https://doi.org/10.1136/bmjopen-2023-073783

Vid Bijelić[a*], Beth Potter[a], Patricia C. Parkin[b], Franco Momoli[a], Mira Liebman[c], Jemila S. Hamid[d]

[a]*University of Ottawa Faculty of Medicine, School of Epidemiology and Public Health, Ottawa, Ontario, Canada;* [b]*Hospital for Sick Children, Pediatrics, Toronto, Ontario, Canada;* [c]*University of Ottawa Faculty of Medicine, Department of Medicine, Ottawa, Ontario, Canada;* [d]*University of Ottawa, Department of Mathematics and Statistics, Ottawa, Ontario, Canada.*

*Corresponding author:
Vid Bijelić. University of Ottawa Faculty of Medicine, School of Epidemiology and Public Health, 600 Peter Morand Crescent, Ottawa, Ontario, Canada K1G 5Z3. vbije016@uottawa.ca. +1 613-276-1487.

Abstract

Introduction: Reference intervals and reference curves provide clinicians with a point of reference when evaluating patients' laboratory test results. In practical applications, the 2.5th and 97.5th percentiles of healthy reference population are typically used as lower and upper reference limits. Guidelines outlining analytical and methodological steps involved in reference intervals and curves estimation are available and there have been large-scale world-wide initiatives to provide reference intervals and curves for children. However, there is a lack of synthesized evidence regarding the results of such initiatives in general, but specifically in iron-related biomarkers, ferritin (in serum and plasma) and hemoglobin. Objectives of this review are to identify studies that have produced reference intervals and curves for ferritin and hemoglobin in pediatric populations and to synthesize all available evidence. We also aim to quantify heterogeneity across reference intervals and curves and identify and elucidate sources of heterogeneity, including heterogeneity in the methods employed in their development.

Methods and analysis: Using a comprehensive search strategy, we will identify eligible studies. Following electronic databases will be searched from inception: EMBASE, MEDLINE, SCOPUS, and The Cochrane Library. We will also perform grey literature search to capture unpublished reference intervals and curves from healthy cohorts. Two researchers will independently screen retrieved citations against eligibility criteria in two stages, focusing first on titles and abstracts and then on full-text articles. Studies that provide reference intervals and curves for ferritin and hemoglobin for pediatric population will be eligible. Data extraction will include study characteristics, characteristics of reference population, methodological and analytical considerations and estimated reference intervals and curves. We will consider narrative synthesis and quantitative synthesis when appropriate.

Ethics and dissemination: Ethical approval is not required as data from already published studies will be used. Results will be disseminated through peer-reviewed publications and conference presentations.

PROSPERO registration number: CRD42023399802

Strengths and limitations of this study

- This is the first systematic review that incorporates quantitative synthesis of pediatric reference intervals and reference curves for ferritin (in serum and plasma) and hemoglobin.
- This review will address inconsistency in estimating reference intervals and reference curves for ferritin and hemoglobin for children of the same age.
- We will quantify heterogeneity across studies as well as identify and elucidate potential sources of heterogeneity.
- This study is limited to articles in English language only.

Introduction

Reference intervals (RIs) and reference curves (RCs) provide clinicians with a point of reference when evaluating patients' laboratory test results [1–3]. They are estimated using a sample of individuals from a healthy population and are often presented as lower (often 2.5th percentile) and upper (often 97.5th percentile) limits, although there are scenarios where only lower or upper limit is used [1,4–6]. With RIs, separate RIs are often required for different sub-groups of the population (for example, by sex or age groups), which can be achieved through partitioning [1]. RCs consist of a series of smooth curves which represent centiles of the distributions of the measurements with respect to age [7–9]. With RCs, age is modeled as a continuous variable, avoiding the need for age partitioning [7–9]. For some analytes, separate RCs may be needed for different subgroups of the population as well (e.g. male vs female).

Estimating RIs and RCs is particularly challenging for pediatric populations. Over the course of childhood, rapid growth and physiological change occurs, necessitating finer age strata than for adult populations [1,10,11]. In addition to the traditional partitioning factors used in adults (e.g., age, sex), other partitioning variables may be needed for pediatric populations including maturity markers, such as Tanner stage, because RIs and RCs for certain biomarkers may vary according to developmental stages [10,12,13]. Achieving adequate enrollment in pediatric populations can be challenging as blood procurement from small children can be difficult. For example, 10ml of blood could constitute 10% of blood volume for a small child [14]. This leads to very sparse data and sample sizes that may be inadequate to provide accurate and reliable RI and RC estimates [6].

While the last several decades have seen large-scale efforts to establish high-quality RIs for children, there is a scarcity of synthesized evidence regarding available RIs for common pediatric biomarkers. Among the most common pediatric laboratory tests are biomarkers associated with iron deficiency (ID), where measurement of hemoglobin (Hb) concentration is used for the assessment of anemia and ferritin (in serum and plasma) concentrations for the assessment of iron status [15]. Iron deficiency is defined as an iron status insufficient to maintain normal physiological functioning [15,16]. It affects about 42% of children worldwide [17].

Analytical factors are also important in reliable estimation of RIs and RCs. Standardization efforts have been undertaken to ensure consistency in measurements of ferritin (in serum and plasma) and Hb concentration. In 1966, haemiglobincyanide (HiCN) spectrophotometric determination was accepted as the international method for determining Hb concentration in human blood [18,19]. Since its endorsement in 1966, the HiCN recommendations have been subjected to regular revisions [20–22]. To address comparability of ferritin results among different assays, the first international standardized reference material (SRM) for serum immunoassay was adopted in 1985 (IS 80/602) [19,23]. This allowed the tracing of the readings of a measuring instrument to a known international reference standard [19,24]. Three more SRM for plasma and serum ferritin were released since 1985 [24,25]. It has been shown that heterogeneity in ferritin (in serum and plasma) estimates exists between different SRM [24].

To our knowledge an up to date and comprehensive systematic assessment and synthesis of pediatric RIs and RCs for ferritin (in serum and plasma) and Hb do not exist. A previous narrative review focused on variability of Hb across the life cycle and published in 2017 [26] did not include a meta-analysis. A number of pediatric RIs and RCs for Hb have been published since 2017. A more recent narrative review of ferritin (in serum and plasma) and Hb RIs also did

not incorporate a meta-analysis and focused only on studies published between 2015 and 2021 retrieved from a search in a single electronic database, PubMed [27].

Objectives

The main objectives of this study are to identify studies that have produced RIs and RCs for ferritin (in serum and plasma) and Hb in the pediatric population and to synthesize all available evidence. We also aim to quantify heterogeneity across studies as well as identify and elucidate potential sources of heterogeneity, including heterogeneity in statistical and analytical methods used in their development.

Materials and Methods

Writing of this protocol for a systematic review and meta-analysis is facilitated and informed by the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol (PRISMA-P) guidelines [28]. The study protocol is registered with PROSPERO (registration number: CRD42023399802). After approval of the protocol, any important amendments to this protocol will be documented in the final publication. If necessary, these changes will be registered with PROSPERO.

Eligibility criteria

Studies that provide RIs or RCs for ferritin (in serum and plasma) and Hb are eligible. Studies that focus on cut-offs, thresholds, or decision limits, but do not provide RI or RC will be excluded. The population of interest is defined as a healthy reference population between birth and 18 years of age. According to CLSI guideline “Health is a relative condition lacking a universal definition.” and term "healthy” therefore becomes the problem in any study [1]. In this systematic review, we follow the CLSI guideline and include studies that, at a minimum, evaluate the health status of reference population using questionnaire and endorse direct sampling techniques. As such, we will include children without acute or chronic disease, not

hospitalized or attending a specialized clinic. Reference intervals or curves established based on samples from hospitalized patients requiring indirect sampling techniques will not be included. Studies including both pediatric and adult populations will be included only if a pediatric population can be separated from the adult population. Due to lack of expertise and resources for translation from other languages the article search will be limited to English language articles. Excluded documents based solely on language will be tracked and the feasibility of their translation will be evaluated, which will depend on their number, potential importance, and funds available. The systematic review for ferritin (in serum and plasma) and Hb will be conducted from inception until July 31, 2023.

Data sources and search strategy

A comprehensive literature search will be conducted on the following electronic databases: EMBASE, MEDLINE, SCOPUS, and The Cochrane Library. Three concepts were pre-identified to inform the search criteria. These concepts are: “pediatric”, “RIs/RCs”, and “hemoglobin/ferritin”. Within each concept, a list of MeSH headings and keywords or phrases was developed. Search terms within each concept were combined with “OR”, and concepts were combined with “AND”. The search strategy for EMBASE is shown in supplemental material section. In addition, we will scan the reference lists of included studies to obtain studies not captured by our literature search, as well as perform grey literature search for any additional non-published (in scientific journals) RIs or RCs. This involves searching some of the well-established pediatric studies and cohorts related to RIs and RCs. In situations where the RIs are not published and are retrieved from websites, we will contact the authors for more information. We will consider studies from Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER), the German Health Interview and Examination Survey for Children and Adolescents (KiGGS), Australian Harmonising Age Pathology Parameters in Kids! (HAPPI

Kids!), Scandinavian Initiative for the Establishment of Pediatric Reference Intervals (NORICHILD), Lifestyle Of Our Kids! (LOOK), and any other major study or cohort our literature review reveals. Our electronic search strategy, shown in supplemental material section, was developed in collaboration with an experienced research librarian from the University of Ottawa.

Study screening and selection

After removing duplicates, two reviewers will independently screen titles and abstracts of retrieved citations. Those citations that are deemed eligible or potentially eligible (those that cannot be ruled out) will move on to the second stage. Discrepancies will be resolved by discussion. For stage 2, two reviewers will independently screen the full texts of articles that passed the first stage. Prior to screening full text results, a pilot test of screening with 10 articles will be conducted to ensure reliability of inclusion/exclusion amongst reviewers. Agreement between reviewers will be assessed using the Kappa statistics. Kappa values will be calculated and strong agreement with values > 0.8 will be sought [29]. Discrepancies between reviewers will be discussed and, if necessary, mediated by a third reviewer. We will use the Covidence software tool for screening [30].

Data extraction

Data will be extracted according to a standardized data abstraction form. To ensure consistency in data extraction, the data extraction form will be piloted on 5 studies by two independent extractors. The extraction form will be modified based on the pilot data if needed (for example, to clarify fields). The data extraction form will include study characteristics (e.g., publication year, country in which the study was conducted), characteristics of the reference population (e.g., type and source, age and sex composition, any other covariates or partitioning variables related to the reference population), methodological characteristics (e.g., statistical methods used to

estimate RIs and RCs, how outliers were handled, partitioning variables, how partitioning was done etc.), analytical characteristics of RIs and RCs (upper and lower limits for each partition for RIs with the corresponding CIs (e.g., 90%), sample size for each partition, partition ranges for continuous variables, regression equation and parameters for RCs), and characteristics of analyte measuring systems (e.g., machine manufacturing and stated traceability to SRM). We will also extract any iron-related factors (e.g., prematurity and birthweight) used for excluding/including children from the studies, as well as any relevant data related to regional and sub-population differences. As values will be extracted for each partition from individual studies, each study will be given a unique identifier to facilitate data management. The data extraction process will be completed by the first reviewer. The extracted data will be verified by a second reviewer.

Risk of bias (RoB) assessment

The RoB of the included studies will be evaluated using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) [31] tailored toward RIs and RCs estimation. We will also consider CLSI guidelines and its recommendations for establishing RIs. In addition, we will consider methodological aspects of RI estimation based on extensive simulations and empirical evaluations published by Daly and colleagues [6]. Daly's paper compared performance of CLSI recommended methods for various sample sizes, variabilities, and levels of skewness and provides additional recommendations in terms of optimality of the methods, in particular when establishing RIs for analytes/biomarkers with skewed distributions. Once relevant RoB components are identified and tailored toward RI and RC estimation, two authors will independently assess each included study against the RoB components. Disagreements will be resolved through discussions and/or by consensus with the second reviewer and other team members. The quality of the studies will be presented in descriptive table(s).

Data synthesis and analysis

The results from the systematic review will first be summarized descriptively, where study characteristics, population characteristics, analytical and statistical methods as well as any relevant data will be summarized and reported. We will consider narrative synthesis as well as quantitative synthesis, when appropriate. Statistical heterogeneity of the reference limits will be evaluated using the I^2 statistic [32], where the lower and upper reference limits are considered separately. We expect significant heterogeneity across the reference limits estimates due to differences in methodological and analytical considerations (e.g., analyte measuring systems) as well as population and regional characteristics across studies. We will examine the extent of heterogeneity and aim to identify and elucidate sources of heterogeneity. Where appropriate, we will conduct random effects meta-analysis to provide pooled estimates for lower and upper reference limits. If data allows, we will perform meta-regression and perform subgroup analysis [32].

In our quantitative synthesis and meta-analysis, we will consider 1-year partitions for children older than 3 years of age and 3 months partitions for those younger than 3 years of age, with 15 days partition for ferritin (in serum and plasma) because of previous studies indicating the dynamic nature of these biomarkers in early life [33]. Considering that there are no standardized age partitions in reference interval literature, we feel one of the contributions of our systematic review is to address this particular challenge involving partitions. As such, we will carefully review studies for possible guidance and recommendations, and our age partitions will be guided by those recommendations. For instance, we will consider studies by Parkin et. al. (2017) and Hamid et. al. (2021) [33,34]. Age partitions will be limited to available data from the individual studies included in our systematic review.

Our literature search will include all studies that reported RIs and RCs for ferritin (in serum and plasma) and Hb. However, some of the reference intervals may involve measurements performed using outdated technology and/or the studies might have used outdated analytical methodology. This information will be gathered during data abstraction/collection and necessary sensitivity analyses will be performed. Sensitivity analysis will also be performed to understand the effect of assay harmonization activities on the pooled results as well as in our risk of bias assessment.

Publication bias and confidence in cumulative evidence

Our systematic review will synthesize studies that estimate RIs and RCs, and not those that test hypothesized associations between variables. As such, publication bias, which is concerned about the possibility of negative or null results not being published [35,36], is not relevant in our systematic review.

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) adaptations exist for studies that do not incorporate comparative stud, for example the GRADE adaptation for prognostic studies [37]. However, to our knowledge, such an adaptation does not exist for studies estimating RIs and/or RCs, and hence no specific tool for assessment of confidence in cumulative evidence for such studies exists. Given that the GRADE handbook [38] discourages researchers from using "modified" GRADE approaches, we will describe confidence in cumulative evidence narratively. This will be done in consultation with clinicians, statisticians, and epidemiologists on our team, and in with consideration of recommendations in the CLSI guidelines and the results of extensive simulations and methodological evaluation conducted by Daly and colleagues [6].

Patient and public involvement

The immediate knowledge users for this systematic review will be clinicians who use RIs and RCs and developers of RIs and RCs, and we have individuals from both of those communities involved as part of the research team. Patients, family members, and the public are not direct knowledge users of the information that will be generated from this systematic review, so we have not engaged them at this stage of our research.

Ethics and Dissemination

Ethical approval is not required as data from already published studies will be used. Results will be disseminated through peer-reviewed publications and conference presentations.

Discussion

The goal of this systematic review is to perform a comprehensive review and to synthesize all available evidence related to pediatric RIs and RCs for ferritin (in serum and plasma) and Hb. We will quantify heterogeneity across RIs and RCs as well as identify and elucidate sources of heterogeneity, including what data sources, types and methods were used. To minimize the risk of missing key citations, we will include the SCOPUS broad subject area database in our search, hand search for relevant studies in references of systematic reviews and primary studies, as well as solicit articles from Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER), the German Health Interview and Examination Survey for Children and Adolescents (KiGGS), Australian Harmonising Age Pathology Parameters in Kids! (HAPPI Kids!), Scandinavian Initiative for the Establishment of Pediatric Reference Intervals (NORICHILD), Lifestyle Of Our Kids! (LOOK), and any other major study or cohort our literature review reveals.

Results from this systematic review will inform the assessment of iron deficiency in children, which is associated with increased morbidity and mortality and affects millions of

children worldwide [39,17]. Indirectly, results from this review will help in our understanding of the broader issues that surround pediatric laboratory RIs and RCs, including providing a better understanding of the methods that have been commonly used and the risk of bias associated with studies publishing RIs and RCs. This systematic review will be limited to articles published in English language.

Acknowledgements
We would also like to thank librarians Karine Fournier and Peter Farrell at the University of Ottawa for their invaluable help in development of the search strategy.

Contributors
VB, JSH, and BP conceptualised and designed the protocol. VB drafted the initial manuscript. VB, JSH, and BP planned the data extraction and statistical analysis, defined the search items, as well as methodological appraisal of the studies. PCP, FM, and ML provided critical insights. All authors have reviewed, approved, and contributed to the final manuscript.

Funding
JSH's research is supported by the Natural Sciences and Engineering Research Council of Canada, Government of Canada (Grant number: RGPIN-2018-06693).

Role of sponsor or funder
The funder had no role in the design and development of this protocol.

Competing interests
None declared.

References

1 Clinical and Laboratory Standards Institute (CLSI), editor. Defining, Establishing and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition. Wayne, Pa: Clinical and Laboratory Standards Institute, 2008.

2 Ceriotti F, Hinzmann R, Panteghini M. Reference intervals: the way forward. *Ann Clin Biochem Int J Lab Med* 2009; **46**: 8–17.

3 Whyte MB, Kelly P. The normal range: it is not normal and it is not a range. *Postgrad Med J* 2018; **94**: 613–6.

4 Higgins V, Truong D, White-Al Habeeb NMA, Fung AWS, Hoffman B, Adeli K. Pediatric reference intervals for 1,25-dihydroxyvitamin D using the DiaSorin LIAISON XL assay in the healthy CALIPER cohort. *Clin Chem Lab Med CCLM* 2018; **56**: 964–72.

5 Aldrimer M, Ridefelt P, Rödöö P, Niklasson F, Gustafsson J, Hellberg D. Population-based pediatric reference intervals for hematology, iron and transferrin. *Scand J Clin Lab Invest* 2013; **73**: 253–61.

6 Daly CH, Higgins V, Adeli K, Grey VL, Hamid JS. Reference interval estimation: Methodological comparison using extensive simulations and empirical data. *Clin Biochem* 2017; **50**: 1145–58.

7 Loh TP, Antoniou G, Baghurst P, Metz MP. Development of paediatric biochemistry centile charts as a complement to laboratory reference intervals. *Pathology (Phila)* 2014; **46**: 336–43.

8 Higgins V, Adeli K. Advances in pediatric reference intervals: from discrete to continuous. *J Lab Precis Med* 2018; **3**: 3–3.

9 Asgari S, Higgins V, McCudden C, Adeli K. Continuous reference intervals for 38 biochemical markers in healthy children and adolescents: Comparisons to traditionally partitioned reference intervals. *Clin Biochem* 2019; **73**: 82–9.

10 Daly CH, Liu X, Grey VL, Hamid JS. A systematic review of statistical methods used in constructing pediatric reference intervals. *Clin Biochem* 2013; **46**: 1220–7.

11 Adeli K, Higgins V, Trajcevski K, White-Al Habeeb N. The Canadian laboratory initiative on pediatric reference intervals: A CALIPER white paper. *Crit Rev Clin Lab Sci* 2017; **54**: 358–413.

12 Huang Y, Eapen E, Steele S, Grey V. Establishment of reference intervals for bone markers in children and adolescents. *Clin Biochem* 2011; **44**: 771–8.

13 Konforte D, Shea JL, Kyriakopoulou L, *et al.* Complex Biological Pattern of Fertility Hormones in Children and Adolescents: A Study of Healthy Children from the CALIPER Cohort and Establishment of Pediatric Reference Intervals. *Clin Chem* 2013; **59**: 1215–27.

14 Clejan S. Neonatology & Laboratory Medicine. Anne Green, Imogen Morgan, and Jim Gray. London, UK: ACB Venture Publications, 2003, 328 pp., £24. ISBN 0-902429-41-8. *Clin Chem* 2004; **50**: 460–460.

15 Baker RD, Greer FR, The Committee on Nutrition. Diagnosis and Prevention of Iron Deficiency and Iron-Deficiency Anemia in Infants and Young Children (0–3 Years of Age). *Pediatrics* 2010; **126**: 1040–50.

16 Abdullah K, Thorpe KE, Maguire JL, *et al.* Risk factors, practice variation and hematological outcomes of children identified with non-anemic iron deficiency following screening in primary care setting. *Paediatr Child Health* 2015; **20**: 302–6.

17 World Health Organization. WHO guidance helps detect iron deficiency and protect brain development. 2020. https://www.who.int/news/item/20-04-2020-who-guidance-helps-detect-iron-deficiency-and-protect-brain-development (accessed July 17, 2023).

18 International Committee for Standardization in Haematology. Recommendations for haemoglobinometry in human blood. *Br J Haematol* 1967; **13**: 71–5.

19 Thorpe SJ. The development and role of international biological reference materials in the diagnosis of anaemia. *Biologicals* 2010; **38**: 449–58.

20 International Committee for Standardization in Haematology. Recommendations for reference method for haemoglobinometry in human blood (ICSH standard EP 6/2: 1977) and specifications for international haemiglobincyanide reference preparation (ICSH standard EP 6/3: 1977). *J Clin Pathol* 1978; **31**: 139–43.

21 Verwilghen RL. Recommendations for reference method for haemoglobinometry in human blood (ICSH standard 1986) and specifications for international haemiglobincyanide reference preparation (3rd edition): INTERNATIONAL COMMITTEE FOR STANDARDIZATION IN HAEMATOLOGY; EXPERT PANEL ON HAEMOGLOBINOMETRY. *Clin Lab Haematol* 1987; **9**: 73–9.

22 Zwart A, Van Assendelft OW, Bull BS, England JM, Lewis SM, Zijlstra WG. Recommendations for reference method for haemoglobinometry in human blood (ICSH standard 1995) and specifications for international haemiglobinocyanide standard (4th edition). *J Clin Pathol* 1996; **49**: 271–4.

23 Fielding CJ, Bothwell TH, Cook JD, *et al.* Preparation, characterization and storage of human ferritin for use as a standard for the assay of serum ferritin. *Clin Lab Haematol* 1984; **6**: 177–87.

24 Braga F, Pasqualetti S, Frusciante E, Borrillo F, Chibireva M, Panteghini M. Harmonization Status of Serum Ferritin Measurements and Implications for Use as Marker of Iron-Related Disorders. *Clin Chem* 2022; **68**: 1202–10.

25 Fox B, Roberts G, Atkinson E, Rigsby P, Ball C. International collaborative study to evaluate and calibrate two recombinant L chain Ferritin preparations for use as a WHO International Standard. *Clin Chem Lab Med CCLM* 2022; **60**: 370–8.

26 Jorgensen JM, Crespo-Bellido M, Dewey KG. Variation in hemoglobin across the life cycle and between males and females. *Ann N Y Acad Sci* 2019; : nyas.14096.

27 Lyle AN, Pokuah F, Dietzen DJ, *et al.* Current State of Pediatric Reference Intervals and the Importance of Correctly Describing the Biochemistry of Child Development: A Review. *JAMA Pediatr* 2022; **176**: 699.

28 PRISMA-P Group, Moher D, Shamseer L, *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015; **4**: 1.

29 Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971; **76**: 378–82.

30 Veritas Health Innovation. Covidence systematic review software. 2022; published online June 28.

31 Whiting PF. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011; **155**: 529.

32 Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21**: 1539–58.

33 Parkin PC, Hamid J, Borkhoff CM, *et al.* Laboratory reference intervals in the assessment of iron status in young children. *BMJ Paediatr Open* 2017; **1**: e000074.

34 Hamid JS, Atenafu EG, Borkhoff CM, *et al.* Reference intervals for hemoglobin and mean corpuscular volume in an ethnically diverse community sample of Canadian children 2 to 36 months. *BMC Pediatr* 2021; **21**: 241.

35 Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629–34.

36 Lin L, Chu H. Quantifying publication bias in meta-analysis: Quantifying Publication Bias. *Biometrics* 2018; **74**: 785–94.

37 Iorio A, Spencer FA, Falavigna M, *et al.* Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients. *BMJ* 2015; **350**: h870–h870.

38 Holger Schünemann, Jan Brożek, Gordon Guyatt, Andrew Oxman. GRADE handbook. 2013. https://gdt.gradepro.org/app/handbook/handbook.html#h.f7lc8w9c3nh8 (accessed March 9, 2023).

39 Scott S, Chen-Edinboro L, Caulfield L, Murray-Kolb L. The Impact of Anemia on Child Mortality: An Updated Review. *Nutrients* 2014; **6**: 5915–32.

Supplemental materials

Hemoglobin search strategy - Medline (Ovid) <1946 to March 02, 2022>

1 Reference Values/ 163043

2 Reference interval*.ti,ab,kf. 5961

3 reference value*.ti,ab,kf. 21918

4 reference range*.ti,ab,kf. 12213

5 normal value*.ti,ab,kf. 25139

6 normal range*.ti,ab,kf. 45154

7 reference interval*.ti,ab,kf. 5961

8 reference curve*.ti,ab,kf. 881

9 cutoff.ti,ab,kf. 55747

10 cut-off.ti,ab,kf. 69610

11 or/1-10 367291

12 Hemoglobins/ 71162

13 H?emoglobin*.ti,ab,kf. 186308

14 Hematology/ 7140

15 H?ematology*.ti,ab,kf. 28408

16 or/12-15 236295

17 adolescent/ or exp child/ or exp infant/ 3812167

18 (child* or youth? or (young adj1 (people or person*)) or p?ediatric* or girl? or boy? or adolescen* or teen* or infant* or newborn*).ti,ab,kf. 2445610

19 (school* adj2 age*).ti,ab,kf. 33697

20 or/17-19 4446539

21 11 and 16 and 20 2751

Hemoglobin search strategy - Embase (Ovid) <1974 to 2022 March 03>

1 Reference Values/ 69648

2 Reference interval*.ti,ab,kw. 8658

3 reference value*.ti,ab,kw. 31299

4 reference range*.ti,ab,kw. 19985

5 normal value*.ti,ab,kw. 36338

6 normal range*.ti,ab,kw. 70607

7 reference interval*.ti,ab,kw. 8658

8 reference curve*.ti,ab,kw. 1166

9 cutoff.ti,ab,kw. 84304

10 cut-off.ti,ab,kw. 128822

11 or/1-10404465

12 Hemoglobins/ 176740

13 H?emoglobin*.ti,ab,kw. 261467

14 Hematology/ 93314

15 H?ematology*.ti,ab,kw. 74501

16 or/12-15 457990

17 adolescent/ or exp child/ or exp infant/ 3660545

18 (child* or youth? or (young adj1 (people or person*)) or p?ediatric* or girl? or boy? or adolescen* or teen* or infant* or newborn*).ti,ab,kw. 2985364

19 (school* adj2 age*).ti,ab,kw. 42018

20 or/17-19 4460807

21 11 and 16 and 20 3879

Hemoglobin search strategy - SCOPUS

Query	Documents
(((reference PRE/2 (value* OR range* OR interval* OR curve * OR centile* OR chart*)) OR (cutoff* OR cut-off*)) AND (h?emoglobin* OR (common PRE/2 h?ematolog*)) AND (p?ediatric* OR child* OR youth? OR (young PRE/0 (people OR person*)) OR girl? OR boy? OR adolescen* OR teen* OR infant* OR newborn* OR (school* PRE/2 age*)) AND (PUBYEAR > 2016))	1,994

Serum and Plasma Ferritin search strategy - Medline (Ovid) <1946 to March 03, 2022>

1 Reference Values/ 163049

2 Reference interval*.ti,ab,kf. 5962

3 reference value*.ti,ab,kf. 21921

4 reference range*.ti,ab,kf. 12215

5 normal value*.ti,ab,kf. 25143

6 normal range*.ti,ab,kf. 45155

7 reference interval*.ti,ab,kf. 5962

8 reference curve*.ti,ab,kf. 881

9 cutoff.ti,ab,kf. 55763

10 cut-off.ti,ab,kf. 69625

11 or/1-10367331

12 Ferritins/ 20673

13 Ferritin*.ti,ab,kf. 31839

14 Hematology/ 7142

15 H?ematology*.ti,ab,kf. 28413

16 or/12-15 67550

17 adolescent/ or exp child/ or exp infant/ 3812718

18 (child* or youth? or (young adj1 (people or person*)) or p?ediatric* or girl? or boy? or adolescen* or teen* or infant* or newborn*).ti,ab,kf. 2445979

19 (school* adj2 age*).ti,ab,kf. 33704

20 or/17-19 4447090

21 11 and 16 and 20 1180

Serum and Plasma Ferritin search strategy - Embase (Ovid) <1974 to 2022 March 03>

1 Reference Values/ 69648

2 Reference interval*.ti,ab,kw. 8658

3 reference value*.ti,ab,kw. 31299

4 reference range*.ti,ab,kw. 19985

5 normal value*.ti,ab,kw. 36338

6 normal range*.ti,ab,kw. 70607

7 reference interval*.ti,ab,kw. 8658

8 reference curve*.ti,ab,kw. 1166

9 cutoff.ti,ab,kw. 84304

10 cut-off.ti,ab,kw. 128822

11 or/1-10404465

12 Ferritins/ 51864

13 Ferritin*.ti,ab,kw. 48848

14 Hematology/ 93314

15 H?ematology*.ti,ab,kw. 74501

16 or/12-15 205677

17 adolescent/ or exp child/ or exp infant/ 3660545

18 (child* or youth? or (young adj1 (people or person*)) or p?ediatric* or girl? or boy? or adolescen* or teen* or infant* or newborn*).ti,ab,kw. 2985364

19 (school* adj2 age*).ti,ab,kw. 42018

20 or/17-19 4460807

21 11 and 16 and 20 2118

Serum and Plasma Ferritin search strategy - SCOPUS

Query	Documents
(((reference PRE/2 (value* OR range* OR interval* OR curve * OR centile* OR chart*)) OR (cutoff* OR cut-off*)) AND (ferritin* OR (common PRE/2 h?ematolog*)) AND (p?ediatric* OR child* OR youth? OR (young PRE/0 (people OR person*)) OR girl? OR boy? OR adolescen* OR teen* OR infant* OR newborn* OR (school* PRE/2 age*)))	4,723

CHAPTER 3: PEDIATRIC REFERENCE INTERVALS AND CURVES FOR HEMOGLOBIN ESTIMATED USING DIRECT METHODS: A SYSTEMATIC REVIEW AND META-ANALYSIS

3.1. Preface to Chapter 3

While in Chapter 2 we presented the published protocol for conducting a systematic review and meta-analysis, in Chapter 3 we present the results of the systematic review and meta-analysis of pediatric hemoglobin reference intervals and curves, lower and upper limits, in males and females under 18 years of age. This article also presents an interactive Shiny web application we developed to visualize both the aggregated findings and individual study results (Pediatric Reference Intervals and Curves Evidence Synthesis - Hemoglobin [PRINCES-H]: https://svetric.shinyapps.io/princesH/). To our knowledge, this is the first systematic review of pediatric hemoglobin reference intervals and curves to incorporate meta-analysis, provide a comprehensive synthesis of both lower and upper limits from RIs and RCs, and present the data through an interactive, web-based platform. This chapter corresponds to Manuscript 2 of the thesis, and is published in *International Journal of Laboratory Hematology*.

VB led the study design, screening of references, full text review and selection, data extraction of selected studies, quantitative and narrative data analysis, development of Shiny app, presentation and interpretation of results and has drafted the initial manuscript. MB was involved in screening of references, full text review and selection, data extraction of selected studies, and the review of the manuscript. JL and MP were involved in data extraction from the included studies and revision of the manuscript. BKP, PCP, and JSH were involved in the design, data interpretation, writing and revision of the manuscript. FM and ML were involved in the design, data interpretation and revision of the manuscript.

3.2 Manuscript status: published in *International Journal of Laboratory Hematology*

Pediatric Reference Intervals and Curves for hemoglobin estimated using direct methods: A Systematic Review and Meta-analysis

Bijelić, V., Bijelić, M., Larock, J., Pham, M., Momoli, F., Liebman, M., Potter, B. K., Parkin, P. C., & Hamid, J. S. (2025). Paediatric Reference Intervals and Curves for Haemoglobin Estimated Using Direct Methods: A Systematic Review and Meta-Analysis. *International journal of laboratory hematology*, *47*(4), 588–599. https://doi.org/10.1111/ijlh.14489

Vid Bijelić[1,2], Marijana Bijelić[1], Josh Larock[3], Michael Pham[2,3], Franco Momoli[1], Mira Liebman[2], Beth K. Potter[1], Patricia C. Parkin[4], Jemila S. Hamid[2,3,*]

1 School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
2 Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada
3 Department of Mathematics and Statistics, University of Ottawa, Ottawa, Ontario, Canada
4 The Hospital for Sick Children, Toronto, Ontario, Canada

* Corresponding Author:
Jemila S. Hamid, PhD
STEM Complex, 150 Louis-Pasteur Pvt
University of Ottawa
Ottawa, ON, K1N 6N5, Canada
jhamid@uottawa.ca

Abstract

Introduction: Hemoglobin is a commonly ordered laboratory test, used to assess both individual and population-level health. To interpret test results, laboratories provide reference intervals (RIs) with lower (2.5th%) and upper (97.5th%) limits according to age and sex. Reference curves (RCs) treat age as a continuous variable. The objectives were to synthesize evidence on pediatric hemoglobin RIs/RCs and investigate possible sources of heterogeneity. We placed our findings in the context of the age- and sex-based hemoglobin thresholds to define anemia, recommended for international use by WHO.

Methods: We conducted a systematic review of studies publishing pediatric hemoglobin RIs/RCs (PROSPERO: CRD42023399802). EMBASE, MEDLINE, SCOPUS, and The Cochrane electronic libraries were searched from inception to July 31, 2023. Studies involving unhealthy children, lacking males and females RIs/RCs, or limited to cord-blood RIs/RCs were excluded. Studies adhering to guidelines for RIs development from the Clinical Laboratory Standards Institute (CLSI) and RCs studies reporting confidence intervals (CIs) were included in meta-analysis. Lower and upper male and female RIs limits were pooled for age groups with heterogeneity below 75% I^2. All studies meeting eligibility criteria were included in narrative synthesis. Sources of heterogeneity were analyzed using heatmaps, forest-plots, and Shiny-app.

Results: Of 9123 studies screened 177 were retained for full-text review. We identified 48 eligible studies (63 529 male and 59 969 female participants) from 25 countries (4 continents) published 1938-2023. There was inconsistency in age partitioning and length of age intervals. Meta-analysis was conducted on 13 studies reporting RIs and 2 studies reporting RCs. Pooled estimates for the 0-3 months age group could not be generated for males or females due to paucity of data. For children aged 3 months or older, both lower and upper RI hemoglobin limits generally increased with age, from approximately 100 to 130 g/L and from approximately 130 to

150 g/L, respectively. For visualization of our narrative synthesis of all 48 studies we created a novel web-based computational tool. Sources of heterogeneity included child age, sex, analyzer type, and country. For many studies, the lower RIs were substantially different than WHO anemia thresholds. Study limitations include a small sample size for younger age groups, potentially impacting heterogeneity estimates, reliance on CLSI guidelines due to the lack of a suitable quality assessment tool for RIs/RCs, and restriction to English-language studies. **Conclusion:** Evidence synthesis of locally developed pediatric hemoglobin RIs/RCs revealed substantial heterogeneity, suggesting the need for more rigorously developed estimates that may be used globally along with WHO thresholds to define anemia. Future research is needed on RIs for the youngest children. Percentile curves should be explored to provide continuous hemoglobin charts.

Introduction

Hemoglobin is one of the most commonly ordered laboratory tests used to assess individual and population health status globally[1]. Decreased hemoglobin levels may indicate anemia, a prevalent condition with numerous etiologies that, if untreated, may lead to poor short term and long term health outcomes[2].

Laboratory tests play a crucial role in clinical decision-making for individual patients as well as in population-based surveillance and policy for public health practitioners. Meaningful interpretation of laboratory test results requires an accompanying range of values, calculated from a healthy reference population. These values are commonly referred to as reference intervals (RIs)[3]. It is essential to establish accurate and reliable RIs, which often involves providing specific RIs for different sub-groups (termed partitions), such as age and sex, when appropriate[4]. Interpretation of laboratory results to determine presence of high-risk or disease requires an accompanying single threshold, known as a clinical decision limit (CDL)[3]. CDLs may be derived from clinical outcome studies, guidelines, and expert consensus.

The International Federation of Clinical Chemistry and Laboratory Medicine Committee on Reference Intervals and Decision Limits has developed the concepts of RIs and CDLs in laboratory medicine[3]. The Clinical Laboratory Standards Institute (CLSI) develops laboratory standards[4].

CLSI provides guidelines for establishing RIs, including the selection of an apparently healthy reference population and minimum sample size requirements[4]. Two RI limits (lower and upper) are calculated using recommended statistical methods (parametric, non-parametric, or robust methods). The lower limit represents the 2.5th percentile, and the upper limit represents the 97.5th percentile the distribution of values (with corresponding 90% confidence intervals) in a healthy reference population. Choice of appropriate statistical methods for estimating RIs is

crucial to the accuracy and precision of RIs, and it is important that researchers carefully examine the distribution of the analytes[4]. In addition to the CLSI guideline, recommendations for choosing appropriate methods are available in literature[5]. Reference curves (RCs), which treat age as a continuous variable, avoid the need for extensive age partitioning and address problems arising from insufficient sample sizes and dynamic physiological changes[5]. Nevertheless, their practical application in laboratory and clinical settings is limited. To our knowledge, there is no clinical or methodological guideline (similar to RIs) available for estimating RCs.

Developing RIs for pediatric populations is particularly challenging because many biomarkers show dynamic changes as children grow from birth through adolescence; thus, narrower sex-specific age partitions may be needed to reflect their age, physiological changes, and developmental stages[5]. Many studies struggle to achieve sufficient sample sizes for precise RIs in pediatric populations, especially considering the difficulty of acquiring blood from younger children[5]. Despite efforts over the last several decades to establish high-quality RIs for pediatric biomarkers, they remain highly inconsistent[6].

RIs for hemoglobin should be derived from apparently healthy populations [4]. However, unrecognized nutritional and genetic factors may cause geographic variation, especially affecting the lower reference limit used for anemia assessment[7]. The World Health Organization (WHO) recently updated hemoglobin thresholds to define anemia based on the 5th percentile of the hemoglobin distribution from a pooled international healthy reference sample[1,8]. They applied number of healthy criteria, but insufficient information was available for genetic factors. This single threshold represents a hemoglobin threshold to define anemia. The WHO views this as "an opportunity for global harmonization of hemoglobin thresholds to define anemia across countries, clinical guidelines, and diagnostic laboratories"[8]. This presents a dilemma for

clinicians and laboratories: should they use the lower RI limit from a local reference population or the single WHO threshold? Understanding the global variation in RIs for hemoglobin may help to address this dilemma.

In this study we conducted a systematic review and meta-analysis of the current literature, to identify and synthesize (where possible) RIs and RCs for hemoglobin in healthy pediatric populations. We also examined heterogeneity across studies and identified potential sources of variation in RIs and RCs.

Methods

The protocol for our systematic review has been published and registered with PROSPERO (CRD42023399802)[9]. We reported the results according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines[10].

Search, screening, and data extraction

The search strategy was developed in consultation with a librarian. We searched major medical databases, MEDLINE, EMBASE, SCOPUS, and the Cochrane Library from inception to search date (July 31, 2023). We also searched the reference lists of included studies. We included studies that involved a healthy pediatric population (<18 years), provided RIs or/and RCs for hemoglobin, and were published in English. We excluded studies that did not report separately male or female RIs or RCs, studies reporting only cord-blood RIs or RCs and studies using indirect methods. When multiple studies used the same database, only the study with the largest sample size, or the most recent study if sample sizes were identical, was included in the systematic review. Full search strategy is available in supplemental material section.

Two reviewers (VB, MB) independently performed the first and second stage of screening; discrepancies were resolved by discussion. Two team members (JL, MP) extracted data which was verified by VB. Study details (e.g., year and country), information on age

partitioning and other methods (e.g., the analyzer used, the statistical methods), and the reported lower and upper RI limits for each partition with their corresponding confidence intervals (CIs) was extracted from each study. CIs for the parametric RI method, when not provided, were calculated using the formula provided by Solberg[11]. When the equations representing the RCs were provided, the mid-point of pre-specified age partitions was used to calculate RIs (from the RCs).

Risk of Bias Assessment

There is no established risk of bias (RoB) tool for RIs/RCs, therefore we were unable to perform a formal risk of bias assessment. Instead, we examined whether or not studies adhered to the CLSI guidelines, which is considered a gold standard for estimating RIs. Two reviewers (VB, MB) independently reviewed each study using CLSI criteria for developing RI (outlier detection, calculations/reporting of CIs for RIs, use of the recommended statistical methods for estimation of RIs, and data partitioning). It is worth mentioning that CLSI guidelines were first published in 2008 using Solberg (1987) approved recommendations on RIs[11]. Publication bias, which is concerned about the possibility of negative or null results not being published, is not applicable to this systematic review, which synthesized studies estimating RIs and RCs.

Evidence Synthesis

Data was synthesized using descriptive statistics and meta-analysis. Only RIs from studies that adhered to the CLSI guidelines and RCs that provided CIs were included in the meta-analysis. To overcome the challenges associated with different age partitions being used by different studies, we employed a novel strategy whereby we created standardized age partitions. We used 3-month age ranges for children below 3 years, and 1-year age ranges for children above 3 years. We defined terms “within range” (study age partition corresponded to or was fully embedded

within our standardized age partitions) and "out of range" (study age partition was wider than standardized age partitions) to indicate the extent to which a study's age partition corresponded to our standardized age partition. We used this information to study age partitioning as a source of heterogeneity in the evidence synthesis of hemoglobin RIs and RCs. For each age partition, we performed meta-analyses separately for males and females and for lower and upper RI limits. Heterogeneity was studied graphically using forest plots, using the I^2 statistic and displayed on a heatmap. Pooled estimates of hemoglobin RIs were calculated only when $I^2 \leq 75\%$, as recommended by the Cochrane Handbook for Systematic Reviews. Random effects meta-analysis was used to account for heterogeneity.

Due to high heterogeneity in methods and results across studies, to complement and extend the meta-analysis, we produced a narrative synthesis of all reviewed studies, inclusive of those for which meta-analysis was not possible. We examined the distribution (using the 10^{th} and 90^{th} percentile) of the lower and upper RI limits across all included studies and provided comparative evaluations of the lower RI limits in relation to the recently published age-specific WHO thresholds for hemoglobin[1]. To facilitate visual evaluation of heterogeneity with respect to the many factors our systematic review revealed, including regional and methodological differences, and to overcome the challenge we faced in having too many partitions, we developed a novel web-based graphical and computational tool (Pediatric Reference Intervals and Curves Evidence Synthesis - Hemoglobin [PRINCES-H] tool). The web-based tool allows investigation of heterogeneity of RIs by selecting specific age groups either using WHO or standardized age partitions, choosing different geographic regions (Asia, Africa, North America, Europe) or choosing if study adhered to CLSI guidelines. It provides dynamic visualizations, stratified by age, sex, region, and adherence to the CLSI guidelines, together with detailed

numerical summaries, including medians, interquartile ranges (IQRs), and CIs. The tool highlights sex-specific differences, particularly during critical growth periods such as adolescence, and enables comparisons of RIs across multiple studies to assess variability in reporting. Additionally, users can input specific hemoglobin values for benchmarking against reported RIs. This tool serves as a practical resource for visualizing and analyzing heterogeneity of hemoglobin RIs, facilitating understandings of regional and methodological differences.

All analyses were performed using the R statistical software version 4.3.1; the *metagen* function from *meta* package version 7.0-0, *metafor* package, and the R Shiny App were used for meta-analysis and developing the web-based tool[12].

Patient and public involvement

No patients or members of the public were involved in this systematic review.

Deviations from the published protocol

In the development of our protocol, we planned to use the Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies to assess RoB. However, after further evaluation the tool was deemed not suitable for RoB assessment of RIs and RCs. As per our protocol, we considered using both fixed and random effects meta-analysis. However, we encountered high level heterogeneity across the studies, hence we decided to present (and interpret) the pooled estimates from the random effects models, as it accounts for heterogeneity. Nevertheless, in the forest plots, we still provided pooled estimates from both the fixed effects and the random effects models, mainly to abide with our protocol for completeness of reporting.

Results

The search yielded 14897 citations of which 5774 were duplicates. A total of 9123 titles and abstracts were screened at the first stage, 177 studies were retained for full-text review, and 48

studies[13–60] met all eligibility criteria and were included in our final systematic review (Figure 1, Table S1).

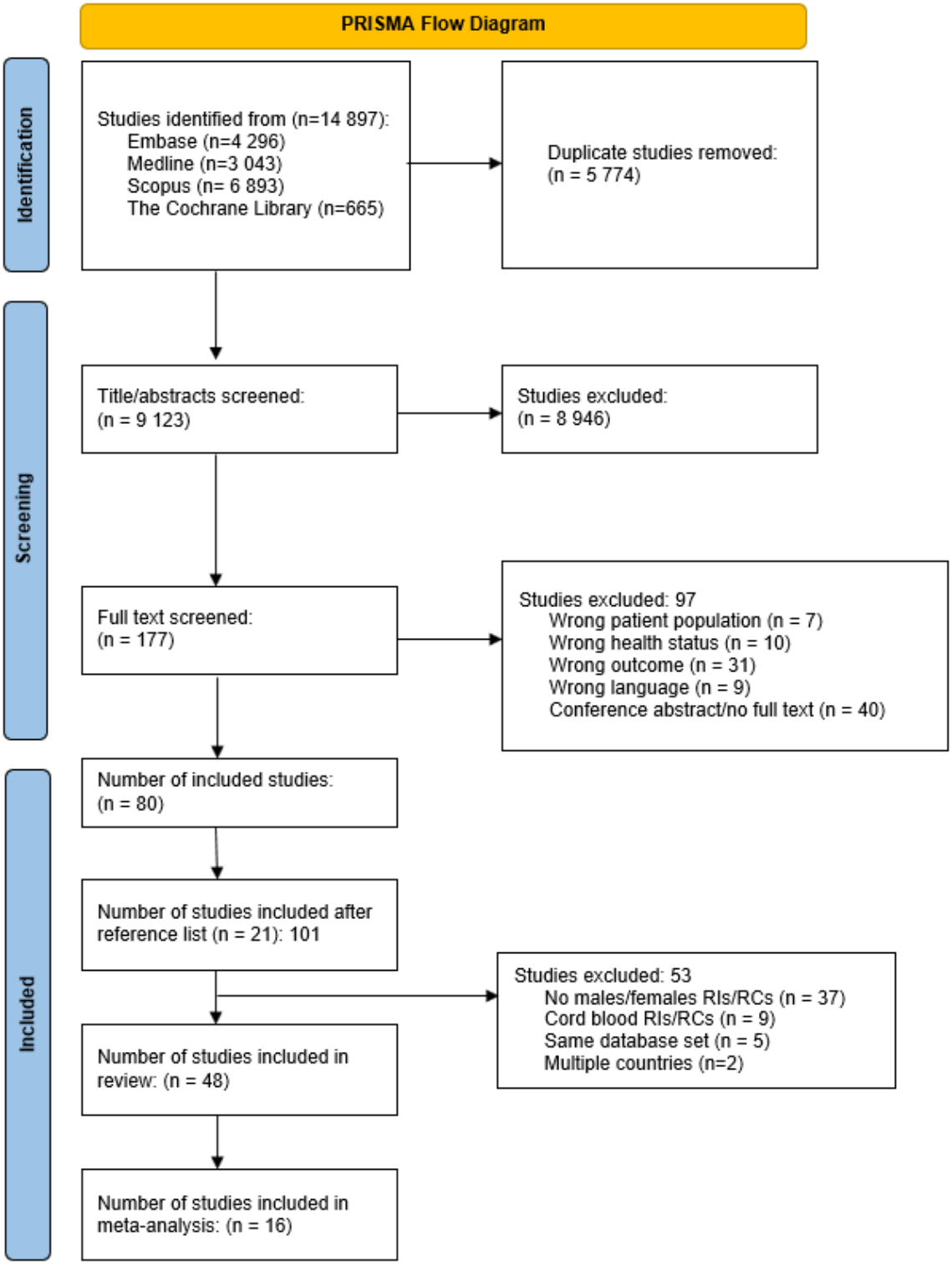


Figure 1. PRISMA flow diagram outlining results from search and screening process. (Adapted from: Page et al. 2020)

Estimates for the lower and upper limits for RIs from 13 studies[13–15,18,24,25,27,29,32,33,36,38,39] that adhered to CLSI guidelines and 2 studies reporting RCs[16,22] that provided CIs for lower and upper limits were considered in the meta-analysis (Table S2) and pooled when appropriate. All 48 studies were included in the narrative synthesis as well as in the graphical examination of heterogeneity.

Table 1. Characteristics of included studies, analytical and methodological characteristics of reference intervals (RIs) and curves (RCs), and study adherence to Clinical Laboratory Standards Institute (CLSI) guidelines

Characteristics*	**Number of Studies, n (%)***	**Number of Children (Sample Size)**
Total	48	123 498
Sex		
Male	48 (100)	63 529 (51.4)
Female	47 (97.9)	59 969 (48.6)
Regions		
Africa	14 (29.2)	5 843 (4.7)
Asia	12 (25.0)	66 038 (53.5)
Europe	9 (18.8)	6 410 (5.2)
North America	13 (27.1)	45 207 (36.6)
Interval Type		
Discrete Reference Interval	44 (91.7)	68 361 (55.4)
Reference Curve	4 (8.3)	55 137 (44.6)
Estimation Method		
Reference Intervals		
Non-parametric	21 (47.7)	28 555 (41.8)
Parametric	18 (40.9)	15 269 (22.3)
Robust	5 (11.4)	668 (1.0)
Other	5 (11.4)	23 869 (34.9)
Reference Curves		
Lambda Mu and Sigma Method	3 (75.0)	54641 (99.1)
Quantile Regression	1 (25.0)	496 (0.9)
Analyzer models		
Abbott	3 (6.2)	1 643 (1.3)
Beckman Coulter	12 (25.0)	40 812 (33.0)
Horiba	2 (4.2)	1 091 (0.9)
Mindray	2 (4.2)	780 (0.6)
Ortho	1 (2.1)	404 (0.3)
Sysmex	20 (41.7)	31 535 (25.5)
Other	4 (8.3)	4 063 (3.3)
Not Provided	5 (10.4)	43 170 (35.0)
Adhered to CLSI Guidelines**	13 (29.5)	36 160 (52.9)
Outlier detection	22 (50.0)	40 881 (59.8)
Partitioning	42 (95.5)	68 108 (99.6)
Method for RIs estimation	32 (72.7)	59 772 (87.4)
Estimation of CIs for RIs	32 (72.7)	60 968 (89.2)

** Does not add up to 100%*

*** Only applicable to RIs (n=44 studies)*

The publication year for the included studies ranged from 1938 to 2023. The geographic distribution covered Africa (n=14, 29.2%), Asia (n=12, 25.0%), Europe (n=9, 18.8%), and North America (n=13, 27.1%) (Table 1). The most commonly used analyzers were the Sysmex, Beckman-Coulter, and Abbott models. Forty-four (91.7%) studies provided RIs; of these, 21 (47.7%) used the non-parametric method for estimating RIs, 18 (40.9%) used the parametric method, and 5 (11.4%) used the robust method. Four of the 48 studies (8.3%) provided RCs, of which 3 used the Lambda Mu and Sigma method and one used quantile regression. Additional summary statistics are provided in the supplementary material section (Table S1).

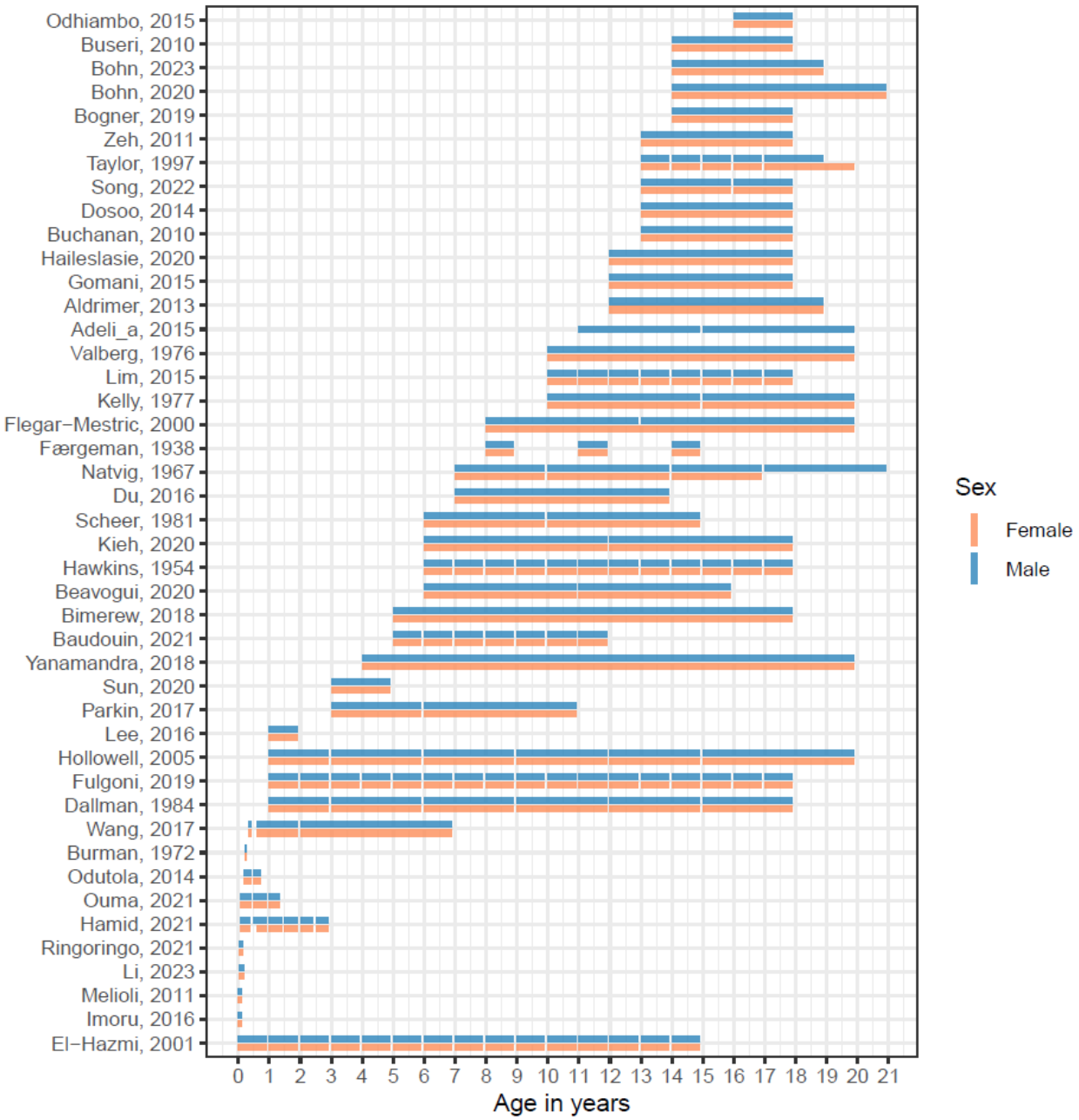


Figure 2. Interval width of the age partitioning used when establishing hemoglobin RIs presented by sex. Studies were ordered (bottom to top) according to the minimum age for the study (reference) population

We observed inconsistencies in age partitioning across studies, for both males and females (Figure 2). The length of age intervals also varied considerably, ranging from less than one month (e.g., birth to 4 days) to as wide as 16 years (e.g., 4 years to 20 years) for both sexes.

Meta-analysis

Sixteen of the 48 (33.3%) studies were included in the meta-analysis. Fourteen RI studies followed all four key recommendations from the CLSI guidelines and two RC studies provided CI for lower and upper limits[4]. The most common reasons for considering a study to be non-adherent to CLSI guidelines were lack of adherence to recommendations related to outlier detection (n=22, 50%) and calculations/reporting of CIs for RIs (n=12, 27.3%), followed by use of other than the recommended statistical methods for estimation of RIs (n=11, 25%), and data partitioning (n=2, 4.5%). One study[36] with missing CIs reported using the parametric method, hence we were able to calculate CIs for this study (Table 1, Table S2). The pooled estimates for the lower and upper limits by age and sex partition, with the corresponding 90% CIs are provided in Table 2. The corresponding forest plots are provided in the supplemental material section (Figures S1-S24).

Table 2. The pooled estimates from random effects meta-analysis for hemoglobin (g/L) reference intervals (RIs).

	Pooled Lower RI Limit (90% CI)		**Pooled Upper RI Limit** (90% CI)	
Age Partitions*	**Males**	**Females**	**Males**	**Females**
0-3 months	**	**	**	**
3-6 months	**	102.3 (98.2-106.4)	133.1 (123.0-143.1)	132.9 (122.8-143.1)
6-9 months	98.6 (94.8-102.4)	**	132.3 (123.6-141.0)	**
9-12 months	100.7 (95.3-106.1)	99.0 (93.5-104.5)	133.1 (120.8-145.3)	**
12-15 months	**	**	**	**
15-18 months	**	**	**	**
18-21 months	103.7 (99.9-107.4)	104.6 (101.2-108.0)	**	**
21-24 months	104.9 (100.7-109.1)	**	**	**
24-27 months	**	**	**	**
27-30 months	**	**	**	**
30-33 months	**	**	**	**
33-36 months	**	**	**	**
3-4 years	**	**	**	**
4-5 years	**	**	**	**
5-6 years	**	**	**	**
6-7 years	113.4 (111.4-115.4)	**	**	**
7-8 years	**	**	**	**
8-9 years	114.4 (112.1-116.7)	**	**	**
9-10 years	**	**	**	**
10-11 years	**	**	**	**
11-12 years	**	116.5 (112.9-120.1)	**	**
12-13 years	122.0 (120.5-123.5)	**	**	**
13-14 years	124.1 (122.8-125.4)	**	**	**
14-15 years	126.0 (124.8-127.3)	**	**	**
15-16 years	128.5 (127.1-130.0)	**	**	**
16-17 years	130.7 (129.1-132.2)	**	**	152.2 (150.6-153.7)
17-18 years	131.7 (130.2-133.3)	**	**	152.2 (150.6-153.8)

** Represents the standardized age-partitions we used in our systematic review*
*** Pooling was not possible because of highly significant heterogeneity I^2>75%*

As shown in Table 2, Figures 3, and Figures S1-S24, we were able to pool the RI estimates for only a limited number of partitions. This was mainly because of the high level of heterogeneity (I^2 > 75%), which is displayed using a heatmap presented in Figure 3. There was a

paucity of data for children aged 0-3 months (Figure S1-S4), and a pooled estimate could not be generated for males or females in this subgroup. Lower limits of the RIs increased with age from 3 months onwards (from approximately 100 g/L to 130 g/L) (Table 2). Upper limits also increased with age from 3 months onwards (from approximately 130 g/L to 150 g/L).

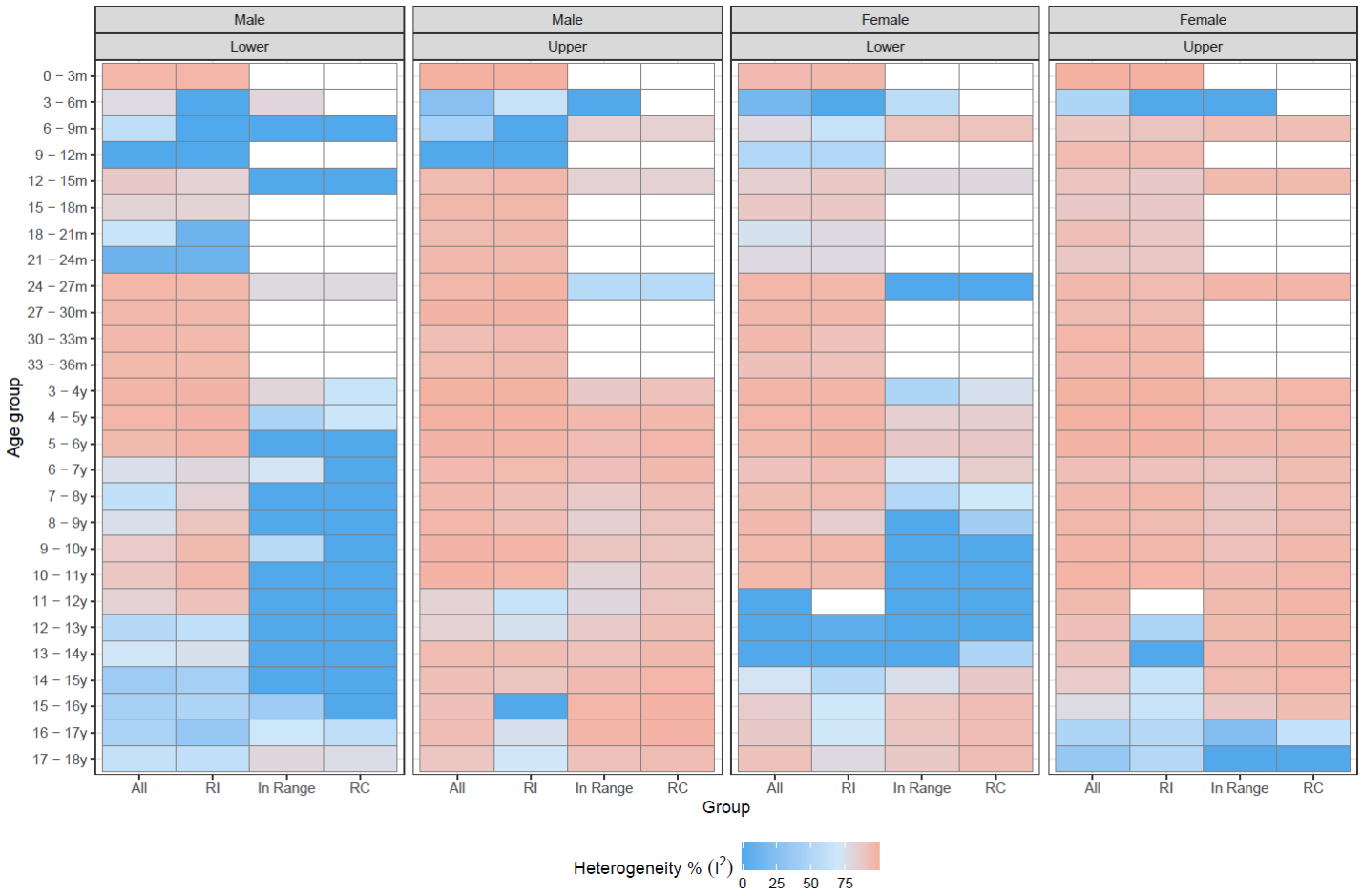


Figure 3. Heatmap representing heterogeneity across age groups and methodological characteristics shown separately for lower and upper limits and males and females. Shown in white are age groups with insufficient sample size to calculate heterogeneity.

In general, lower heterogeneity ($I^2 \leq 75\%$, dark and light blue on the heatmap) was found for males (compared with females), older age partitions with 1-year intervals (compared with younger age partitions with 3-month intervals), narrower age partitions (compared with wider age partitions), "within range" age partitions (compared with "out of range" age partitions), and

lower RI limits (compared with upper RI limits) (Figure 3). When heterogeneity could not be calculated, due to an insufficient number of studies (n<2), cells were left blank on the heatmap.

Narrative Synthesis

Detailed descriptions of the full set of 48 studies are provided in Table S1. The distributions of the RIs for all age partitions, separately for males and females, are presented in the forest plots and descriptive statistics provided in the supplementary files (Figures S1-S24, Table S3), where regional differences are also highlighted. A more comprehensive visualization of the RIs is provided using our novel web-based PRINCES-H tool (https://svetric.shinyapps.io/princesH/) where the user can specify age group, region, and adherence to the CLSI guidelines. The age-specific hemoglobin thresholds published by WHO as well as the 10^{th} and 90^{th} percentiles of RIs across the studies are displayed in figures generated by our tool. The tool also provides a summary narrative synthesis based on the user-specified features. An illustrative example based on one scenario is provided in the supplementary file (Figure S25, Table S3).

Based on the set of 48.0 studies, for children aged 0-6 months, the lower RI limits ranged from 85.0 g/L (males) and 86 g/L (females)[19] to 147.0 g/L (males) and 142.0 g/L (females)[45]. The upper RI limits in this age partition ranged from 117.4 g/L (males) and 116.6 g/L (females)[57] to 224.0 g/L (males) and 227.0 g/L (females)[45] (Figure S26 and S27, Table S3).

For children aged 6-23 months, the lower RI limits ranged from 66.0 g/L (males) and 69 g/L (females)[19] to 112.5 g/L (males) and 122.2 g/L (females)[50]. In this age partition, the upper RI limits ranged from 117.3 g/L (males) and 120.9 g/L (females)[50] to 141.0 g/L (males)[36] and 145.8 g/L (females)[21] (Figure S28 and S29, Table S3).

For children aged 24-59 months, the lower RI limits ranged from 91.1 g/L (males) and 93.0 g/L (females)[21] to 150.8 g/L (males) and 138.0 g/L (females)[31]. The upper RI limits in this

age group ranged from 123.5 g/L (males) and 126.2 g/L (females)[50] to 159.0 g/L (males) and 157.0 g/L (females)[27] (Figure S30 and S31, Table S3).

For children aged 5-12 years, the lower RI limits ranged from 80.9 g/L (males) and 85.9 g/L (females)[23] to 150.8 g/L (males) and 138.0 g/L (females)[31]. The upper RI limits ranged from 117.9 g/L (males) and 117.0 g/L (females)[17] to 196.0 g/L (males)[30] and 162.0 g/L (females)[56] (Figure S32 and S33, Table S3).

For children aged 12-15 years, lower RI limits ranged from 90.0 g/L (males)[23] and 81.0 g/L (females)[46] to 150.8 g/L (males)[31] and 140.5 g/L (females)[58]. The upper RI limits ranged from 133.2 g/L (males) and 133.2 g/L (females)[58] to 196.0 g/L (males)[30] and 162.0 g/L (females)[56] (Figure S34 and S35, Table S3).

For children aged 15-18 years, lower RI limits ranged from 90 g/L (males)[23] and 75.0 g/L (females)[41] to 154.0 g/L (males) and 140.5 g/L (females)[58]. The upper RI limits ranged from 140.5 g/L (males)[23] and 131.5 g/L (females)[59] to 196.0 g/L (males)[30] and 162.0 g/L (females)[56] (Figure S36 and S37, Table S3).

The supplemental Figure S38 highlights the heterogeneity of hemoglobin reference intervals across age groups and between sexes, and the complexity involved in visual presentation of the all the RIs. Since the age ranges for RIs often span months or years, the RI limits and their corresponding CIs were displayed on graphs as point estimates at the midpoint of each age range. Significant variability is observed, particularly during adolescence, where males and females show distinct variation in their reference intervals. Based on the visual presentation of the studies using the PRINCES-H tool, we identified substantial variability related to study characteristics, including child age and sex, analyzer type, and country. There also appeared to

be significant variation by countries within continental regions. For many studies, the lower RI was substantially different (lower or higher) than the WHO threshold to define anemia.

Discussion

We conducted a systematic review of pediatric RIs/RCs for hemoglobin (lower and upper limits) in males and females under 18 years of age. We identified 48 studies from 25 countries spanning 4 continents published between 1938 and 2023. There was inconsistency in age partitioning and length of age intervals across studies. Fourteen studies reporting RIs adhered to CLSI guidelines and were included in the meta-analysis, along with 2 studies that reported RCs. RI estimates were pooled for a limited number of age partitions due to high heterogeneity. A pooled estimate for 0-3 months could not be generated due to a paucity of data. Lower RI limits increased with age from 3 months onwards from approximately 100 g/L to 130 g/L. Upper limits increased from 3 months onwards from approximately 130 g/L to 150 g/L. We also conducted a narrative synthesis of all 48 studies accompanied by a comprehensive visualization of the RIs using our novel web-based PRINCES-H tool, according to our standardized age partitions as well as the WHO age groups. We identified substantial heterogeneity in lower and upper RI limits related to study characteristics, including child age and sex, analyzer type, and country.

To our knowledge, this is the first systematic review on pediatric hemoglobin RIs, yet it has several limitations. First, we did not use a formal RoB or quality assessment because there is no tool applicable for RIs/RCs; instead, we used adherence to the CLSI guideline as a criterion to pool RI estimates. Second, our study was limited to publications in English, which may have led to the exclusion of important studies published in other languages. Third, there was a small number of studies for some of the age groups, especially the younger age groups, which could affect heterogeneity estimates.

While our review focused on RIs/RCs (lower and upper limits), there is another body of literature focused on a single hemoglobin threshold to define anemia (a type of clinical decision limit). Jorgensen et al., commissioned by WHO, conducted a narrative review of 60 studies (1975-2018) reporting hemoglobin thresholds across the life cycle in males and females[7]. Recognizing the substantial heterogeneity in study results and inclusion of individuals with iron deficiency or inflammation, WHO supported a subsequent study by Braat et al. which pooled international data sources and excluded individuals with clinical or laboratory evidence of conditions that might reduce hemoglobin (i.e., creating a healthy reference population)[1]. The authors estimated hemoglobin thresholds at the 5th percentile for five age groups (6 months to 65 years), noting similar thresholds for males and females under 12 years of age[1]. Commenting on the WHO guidelines, Pasricha et al. suggest there is "an opportunity for global harmonization of hemoglobin thresholds to define anemia across countries, clinical guidelines, and diagnostic laboratories"[8]. Our finding of substantial heterogeneity of locally developed RIs/RCs, suggests that global harmonization of RIs is not yet possible, and it remains uncertain whether local or global RIs will lead to improved clinical decision-making. More research on the rigorous development of RIs is warranted. In addition, a discussion of the usefulness and precision of local versus global pediatric RIs is warranted. The question related to the risk of erroneous diagnosis of anemia using global limits for some populations should be investigated in future studies.

Community- or hospital-based data, using indirect methods, play important roles in establishing RIs and RCs. Indirect methods have been explored in several studies as an alternative to direct approaches and have the potential to influence clinical practice, particularly in vulnerable populations such as the pediatric population, where acquiring sufficient sample

sizes from healthy children is often limited or not possible [61–64]. Studies utilizing indirect methods estimate RIs using large databases from hospital information systems and/or community-based datasets, leveraging traditional and emerging statistical techniques including the Hoffmann method and methods involving the mixture distribution [61,65–68]. The CLSI guidelines, which is currently considered as the gold standard for RI estimation, acknowledge the use of indirect techniques. However, the guidelines also highlight reservations, and the recommendations often focus on RIs and RCs established using healthy populations. Considering this and the ongoing debate around direct versus indirect methods, and the anticipated heterogeneity and scale of data in this systematic review, we focused on RIs/RCs derived exclusively from "healthy populations", and only those that followed the CLSI guidelines were combined quantitatively to provide pooled estimates. By identifying and highlighting areas where the estimation of precise RIs is not yet possible, our review provides support for the use of indirect methods using community-based data and hospital samples, as a provisional solution in these areas. At the same time, our findings emphasize the critical importance of ongoing and future pediatric data collection efforts for gathering sufficient samples from healthy children. Our systematic review and meta-analysis employed novel strategies, such as standardized age partitioning, to advance evidence synthesis methodologies specific to RIs/RCs. By limiting the scope of this analysis to "healthy populations", we presented this work as a proof-of-concept for future studies in the field. The methods employed and the analysis strategies used, including our novel web-based tool computational and graphical, can be adapted and used in evidence synthesis involving RIs and RCs for other biomarkers of health, including RIs and RCs obtained using indirect methods.

There are several implications for practice, policy and future research. First, our evidence synthesis of locally developed RIs/RCs revealed substantial heterogeneity, suggesting the need for more rigorously developed estimates that may be used globally, similar to the approach by Braat et al. for development of hemoglobin thresholds to define anemia. Second, standardized age partitions are needed for RIs, and RCs may eliminate the need for age partitioning. Third, future research is needed on RIs for the youngest children, using very narrow age partitions, where hemoglobin changes significantly from birth to 6 months of age. Fourth, we support the aspirational goal of global harmonization of RIs and RCs (lower and upper limits), and hemoglobin thresholds to define anemia (single CDL) which will provide clinicians, public health practitioners and laboratories with a comprehensive set of values to assess the health of individuals and populations. Use of percentile curves should be explored to provide continuous hemoglobin charts (similar to WHO growth charts), which will eliminate the need for age partitioning. Fifth, at the same time we also advocate for regional RIs and RCs to understand the range of values of the population.

ACKNOWLEDGMENT

The authors acknowledge the support from the University of Ottawa in the library search. This research is supported by JSH's grant from the Natural Sciences and Engineering Research Council of Canada, Government of Canada (Grant number: RGPIN-2018-06693). The funder had no role in the design and development of this protocol.

References

1 Braat S, Fielding KL, Han J, *et al.* Haemoglobin thresholds to define anaemia from age 6 months to 65 years: estimates from international data sources. *Lancet Haematol* 2024; **11**: e253–64.

2 Lozoff B, Jimenez E, Smith JB. Double Burden of Iron Deficiency in Infancy and Low Socioeconomic Status: A Longitudinal Analysis of Cognitive Test Scores to Age 19 Years. *Arch Pediatr Adolesc Med* 2006; **160**: 1108.

3 Ozarda Y, Sikaris K, Streichert T, Macri J, on behalf of IFCC Committee on Reference intervals and Decision Limits (C-RIDL). Distinguishing reference intervals and clinical decision limits – A review by the IFCC Committee on Reference Intervals and Decision Limits. *Crit Rev Clin Lab Sci* 2018; **55**: 420–31.

4 Clinical and Laboratory Standards Institute (CLSI), editor. Defining, Establishing and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition. Wayne, Pa: Clinical and Laboratory Standards Institute, 2008.

5 Daly CH, Higgins V, Adeli K, Grey VL, Hamid JS. Reference interval estimation: Methodological comparison using extensive simulations and empirical data. *Clin Biochem* 2017; **50**: 1145–58.

6 Adeli K, Higgins V, Seccombe D, *et al.* National Survey of Adult and Pediatric Reference Intervals in Clinical Laboratories across Canada: A Report of the CSCC Working Group on Reference Interval Harmonization. *Clin Biochem* 2017; **50**: 925–35.

7 Jorgensen JM, Crespo-Bellido M, Dewey KG. Variation in hemoglobin across the life cycle and between males and females. *Ann N Y Acad Sci* 2019; : nyas.14096.

8 Pasricha S-R, Rogers L, Branca F, Garcia-Casal M-N. Measuring haemoglobin concentration to define anaemia: WHO guidelines. *The Lancet* 2024; **403**: 1963–6.

9 Bijelić V, Potter B, Parkin PC, Momoli F, Liebman M, Hamid JS. Paediatric reference intervals and curves for haemoglobin and ferritin: protocol for a systematic review and meta-analysis. *BMJ Open* 2023; **13**: e073783.

10 Page MJ, McKenzie JE, Bossuyt PM, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; **372**: n71.

11 Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *Clin Chim Acta* 1987; **170**: S13–32.

12 R Core Team. R: A language and environment for statistical computing. 2023. https://www.R-project.org/.

13 Bohn MK, Wilson S, Steele S, Adeli K. Comprehensive pediatric reference intervals for 79 hematology markers in the CALIPER cohort of healthy children and adolescents using the Mindray BC-6800PLUS system. *Int J Lab Hematol* 2023; **45**: 469–80.

14 Li C-L, Tan C, Gao L, Chang Y, Fan W-M. Establishment and validation of haematological reference intervals for newborns aged 5 to 28 days in Nanjing, China. *Hematology* 2023; **28**: 2208891.

15 Song W, Yan R, Peng M, *et al.* Age and sex specific reference intervals of 13 hematological analytes in Chinese children and adolescents aged from 28 days up to 20 years: the PRINCE study. *Clin Chem Lab Med CCLM* 2022; **60**: 1250–60.

16 Yan R, Peng Y, Hu L, *et al.* Continuous reference intervals for 21 biochemical and hematological analytes in healthy Chinese children and adolescents: The PRINCE study. *Clin Biochem* 2022; **102**: 9–18.

17 Baudouin KA, Soualio K, Howélé O, Mathieu BN, Paul Y. Variation in blood count parameters of children aged 5 to 11 years in Abidjan, Côte d'Ivoire. *Iran J Pediatr Hematol Oncol* 2021; published online June 29. DOI:10.18502/ijpho.v11i3.6562.

18 Hamid JS, Atenafu EG, Borkhoff CM, *et al.* Reference intervals for hemoglobin and mean corpuscular volume in an ethnically diverse community sample of Canadian children 2 to 36 months. *BMC Pediatr* 2021; **21**: 241.

19 Ouma JO, Mulama DH, Otieno L, *et al.* Clinical laboratory hematology reference values among infants aged 1month to 17 months in Kombewa Sub-County, Kisumu: A cross sectional study of rural population in Western Kenya. *PLOS ONE* 2021; **16**: e0244786.

20 Ringoringo HP. Reference Range of Complete Blood Count in Healthy Term Newborns 1 Week after Birth. *Open Access Maced J Med Sci* 2021; **9**: 1565–9.

21 Sachdev HS, Porwal A, Acharya R, *et al.* Haemoglobin thresholds to define anaemia in a national sample of healthy children and adolescents aged 1–19 years in India: a population-based study. *Lancet Glob Health* 2021; **9**: e822–31.

22 Wilson S, Bohn MK, Hall A, Higgins V, Abdelhaleem M, Adeli K. Continuous reference curves for common hematology markers in the CALIPER cohort of healthy children and adolescents on the Sysmex XN-3000 system. *Int J Lab Hematol* 2021; **43**: 1394–402.

23 Béavogui AH, Touré AA, Camara D, *et al.* Haematological and Biochemical Reference Values for Healthy Population of Maferinyah Rural Community in Guinea. *BioMed Res Int* 2020; **2020**: 1–10.

24 Bohn MK, Higgins V, Tahmasebi H, *et al.* Complex biological patterns of hematology parameters in childhood necessitating age- and sex-specific reference intervals for evidence-based clinical interpretation. *Int J Lab Hematol* 2020; **42**: 750–60.

25 Haileslasie H, Tsegaye A, Teklehaymanot G, *et al.* Community based hematological reference intervals among apparently healthy adolescents aged 12-17 years in Mekelle city, Tigrai, northern Ethiopia: A cross sectional study. *PLOS ONE* 2020; **15**: e0234106.

26 Kieh MW, Browne SM, Grandits GA, *et al.* Adult and paediatric haematology and clinical chemistry laboratory reference limits for Liberia. *Afr J Lab Med* 2020; **9**. DOI:10.4102/ajlm.v9i1.1080.

27 Sun Q, Yue Y, Tian Y, Wang Q. Establishment of reference intervals for pediatric complete capillary blood counts: A multicenter study in Beijing. *Int J Lab Hematol* 2020; **42**: 533–43.

28 Bogner B, Schwenoha K, Vogl M, *et al.* Evaluation of reference intervals of haematological and biochemical markers in an Austrian adolescent study cohort. *Clin Chem Lab Med CCLM* 2019; **57**: 891–900.

29 Fulgoni VL, Agarwal S, Kellogg MD, Lieberman HR. Establishing Pediatric and Adult RBC Reference Intervals With NHANES Data Using Piecewise Regression. *Am J Clin Pathol* 2019; **151**: 128–42.

30 Bimerew LG, Demie T, Eskinder K, *et al.* Reference intervals for hematology test parameters from apparently healthy individuals in southwest Ethiopia. *SAGE Open Med* 2018; **6**: 205031211880762.

31 Yanamandra U, Bhattachar SA, Das SK, *et al.* Hematological Parameters in Native Highlanders of Ladakh Aged 4–19 Years. *Indian J Hematol Blood Transfus* 2018; **34**: 731–8.

32 Parkin PC, Hamid J, Borkhoff CM, *et al.* Laboratory reference intervals in the assessment of iron status in young children. *BMJ Paediatr Open* 2017; **1**: e000074.

33 Wang G-C, Li N, Niu C, *et al.* Establishment of complete blood count reference intervals for Chinese preschoolers. *J Clin Lab Anal* 2017; **31**: e22095.

34 Du Y, Chen H, Sun H, Wang C, Wang C, Li Y. Investigation and Analysis of Reference Intervals for Blood Cell Parameters in a Healthy Population from Daxingan Region Inner Mongolia. *Clin Lab* 2016; **62**. DOI:10.7754/Clin.Lab.2015.150608.

35 Imoru M, Buhari H, Nuhu A, Department of Haematology, Aminu Kano Teaching Hospital, Kano, Nigeria. Haematological Reference Values for Full term Healthy Neonates Delivered within 24 Hours in Sokoto, Northern Nigeria. *Int J Hematol Ther* 2016; **2**. DOI:10.15436/2381-1404.16.013.

36 Lee HR, Shin S, Yoon JH, Roh EY, Chang JY. Reference Intervals of Hematology and Clinical Chemistry Analytes for 1-Year-Old Korean Children. *Ann Lab Med* 2016; **36**: 481–8.

37 Rieger K, Vogel M, Engel C, *et al.* Reference intervals for iron-related blood parameters: results from a population-based cohort study (LIFE Child). *LaboratoriumsMedizin* 2016; **40**: 000010151520160019.

38 Adeli K, Raizman JE, Chen Y, *et al.* Complex Biological Profile of Hematologic Markers across Pediatric, Adult, and Geriatric Ages: Establishment of Robust Pediatric and Adult Reference Intervals on the Basis of the Canadian Health Measures Survey. *Clin Chem* 2015; **61**: 1075–86.

39 Gomani P, Matubu A, Mujuru H, *et al.* Hematological and Biochemical Laboratory Reference Intervals for Zimbabwean Adolescents. *Clin Lab* 2015; **61**. DOI:10.7754/Clin.Lab.2014.140506.

40 Lim JS, Seo J, Kim DH, Lee JA. The hemoglobin values of Korean adolescents show distinctive characteristics in comparison to those of Caucasians and African Americans. *Pediatr Blood Cancer* 2015; **62**: 409–13.

41 Odhiambo C, Oyaro B, Odipo R, *et al.* Evaluation of Locally Established Reference Intervals for Hematology and Biochemistry Parameters in Western Kenya. *PLOS ONE* 2015; **10**: e0123140.

42 Dosoo DK, Amenga-Etego S, Kayan K, *et al.* Biochemical and Hematologic Parameters for Children in the Middle Belt of Ghana. *Am J Trop Med Hyg* 2014; **90**: 767–73.

43 Odutola AA, Afolabi MO, Jafali J, *et al.* Haematological and biochemical reference values of G ambian infants. *Trop Med Int Health* 2014; **19**: 275–83.

44 Aldrimer M, Ridefelt P, Rödöö P, Niklasson F, Gustafsson J, Hellberg D. Population-based pediatric reference intervals for hematology, iron and transferrin. *Scand J Clin Lab Invest* 2013; **73**: 253–61.

45 Melioli G. Reference values of blood cell counts in the first days of life. *Front Biosci* 2011; **E3**: 871–8.

46 Zeh C, Amornkul PN, Inzaule S, *et al.* Population-Based Biochemistry, Immunologic and Hematological Reference Values for Adolescents and Young Adults in a Rural Population in Western Kenya. *PLoS ONE* 2011; **6**: e21040.

47 Buchanan AM, Muro FJ, Gratz J, *et al.* Establishment of haematological and immunological reference values for healthy Tanzanian children in Kilimanjaro Region: Immunohaematological reference ranges for Tanzanian children. *Trop Med Int Health* 2010; : no-no.

48 Buseri FI, Jeremiah Z, Joy S. Reference values of hematological indices of infants, children, and adolescents in Port Harcourt, Nigeria. *Pathol Lab Med Int* 2010; : 65.

49 Hollowell JG, van Assendelft OW, Gunter EW, *et al.* Hematological and iron-related analytes--reference data for persons aged 1 year and over: United States, 1988-94. *Vital Health Stat 11* 2005; : 1–156.

50 El-Hazmi MAF, Warsy AS. Normal Reference Values for Hematological Parameters, Red Cell Indices, Hb A2 and Hb F from Early Childhood Through Adolescence in Saudis. *Ann Saudi Med* 2001; **21**: 165–9.

51 Flegar-Mestrić Z, Nazor A, Jagarinec N. Haematological profile in healthy urban population (8 to 70 years of age). *Coll Antropol* 2000; **24**: 185–96.

52 Taylor MRH, Holland CV, Spencer R, Jackson JF, O'Connor GI, O'Donnell JR. Haematological reference ranges for schoolchildren. *Clin Lab Haematol* 1997; **19**: 1–15.

53 Dallman PR, Yip R, Johnson C. Prevalence and causes of anemia in the United States, 1976 to 1980. *Am J Clin Nutr* 1984; **39**: 437–45.

54 Scheer JC, Guthrie HA. Hemoglobin criteria with respect to obesity. *Am J Clin Nutr* 1981; **34**: 2748–51.

55 Kelly A, Munan L. Haematologic Profile of Natural Populations: Red Cell Parameters. *Br J Haematol* 1977; **35**: 153–60.

56 Valberg LS, Sorbie J, Ludwig J, Pelletier O. Serum ferritin and the iron status of Canadians. *Can Med Assoc J* 1976; **114**: 417–21.

57 Burman D. Haemoglobin Levels in Normal Infants Aged 3 to 24 Months, and the Effect of Iron. *Arch Dis Child* 1972; **47**: 261–71.

58 Natvig H, Vellar OD, Andersen J. Studies on hemoglobin values in Norway. VII. Hemoglobin, hematocrit and MCHC values among boys and girls aged 7-20 years in elementary and grammar schools. *Acta Med Scand* 1967; **182**: 183–91.

59 Hawkins WW, Speck E, Leonard VG. Variation of the hemoglobin level with age and sex. *Blood* 1954; **9**: 999–1007.

60 Færgeman P. The normal values for hemoglobin, erythrocytes and cell volume per cent, and the corpuscular constants derived thereof, in children aged 8–14 years.: Examinations of 300 school-children in Stockholm [1]. *Acta Med Scand* 1938; **95**: 566–96.

61 Hoffmann RG. Statistics in the Practice of Medicine. *JAMA* 1963; **185**: 864–73.

62 Katayev A, Balciza C, Seccombe DW. Establishing reference intervals for clinical laboratory test results: is there a better way? *Am J Clin Pathol* 2010; **133**: 180–6.

63 Concordet D, Geffré A, Braun JP, Trumel C. A new approach for the determination of reference intervals from hospital-based data. *Clin Chim Acta Int J Clin Chem* 2009; **405**: 43–8.

64 Farrell C-JL, Nguyen L. Indirect Reference Intervals: Harnessing the Power of Stored Laboratory Data. *Clin Biochem Rev* 2019; **40**: 99–111.

65 Zierk J, Arzideh F, Haeckel R, Rascher W, Rauh M, Metzler M. Indirect determination of pediatric blood count reference intervals. *Clin Chem Lab Med CCLM* 2013; **51**: 863–72.

66 Zierk J, Hirschmann J, Toddenroth D, *et al.* Next-generation reference intervals for pediatric hematology. *Clin Chem Lab Med CCLM* 2019; **57**: 1595–607.

67 Zeljkovic A, Csuzdi Balog Z, Dukai E, Vekic J, Jelic-Ivanovic Z, Spasojevic-Kalimanovska V. Indirect reference intervals for haematological parameters in capillary blood of pre-school children. *Biochem Medica* 2021; **31**: 134–42.

68 Mrosewski I, Dähn T, Hehde J, *et al.* Indirectly determined hematology reference intervals for pediatric patients in Berlin and Brandenburg. *Clin Chem Lab Med CCLM* 2022; **60**: 408–32.

Supplemental tables and figures

Table S1. Studies included in systematic review

Author	Year	Country	No. of sample	Analyzer	Ref. Type	Age Range	90%CI	CLSI Cited	Daly 2017 Cited	CLSI Followed
Wilson	2023	Canada	52500	Abbott	RC	0 m - 18 y	No	Yes	No	No
Bohn	2021	Canada	165	Siemens	RI	13 y - 19 y	Yes	Yes	No	Yes
Hall	2021	Canada	17800	Not provided	RC	6 m - 18 y	No	Yes	No	Yes
Tahmasebi	2020	Canada	229	Abbott	RI	5 m - 19 m	Yes	Yes	No	Yes
Bogner	2019	Austria	102	Roche	RI	14 y - 18 y	Yes	Yes	No	No
Bohn	2019	Canada	148	Roche	RI	15 y - 19 y	Yes	Yes	No	Yes
Hoq	2019	Australia	51300	Ortho, Abbott, Roche, Siemens, Beckman Coulter	RC	0 m - 18 y	No	Yes	No	No
Larsson	2019	Sweden	826	Roche	RI	0 m - 12 m	Yes	Yes	No	No
Higgins	2018	Canada	197	Ortho	RI	13 y - 19 y	Yes	Yes	No	Yes
Parkin	2017	Canada	5007	Roche	RI	0 m - 11 y	Yes	Yes	No	Yes
Karbasy	2016	Canada	84	Beckman Coulter	RI	16 y - 19 y	Yes	Yes	No	Yes
Rieger	2016	Germany	24906	Roche	RC	4 y - 17 y	No	Yes	No	No
Adeli	2015	Canada	1470	Siemens	RI	6 y - 25 y	Yes	Yes	No	Yes
Bailey	2013	Canada	304	Abbott	RI	14 y - 19 y	Yes	Yes	No	Yes
Southcott	2010	Australia	986	Abbott	RI	8 y - 13 y	No	Yes	No	No
Hollowell	2005	US	10279	Beckman Coulter	RI	12 m - 20 y	Yes	No	No	No
Elmlinger	2002	Germany	742	Siemens	RI	0 m - 18 y	No	Yes	No	No
Chinn	1998	Scotland	753	Other	RI	8 y - 10 y	No	No		No
Wiedermann	1993	Germany	135	Abbott	RI	14 y - 19 y	No	Yes	No	No
Valberg	1976	Canada	204	Not provided	RI	10 y - 20 y	No	No	No	No

Table S2. Study adherence to CLSI guidelines

First Author Publication Year	Outlier detection	Partitioning	Method for RIs estimation	Estimation of CIs for RIs
Valberg et al., (1976)	N	N	Y	Y
Wiedermann et al., (1993)	N	Y	Y	N
Chinn et al., (1998)	N	Y	Y	N
Elmlinger et al., (2002)	N	Y	N	N
Hollowell et al., (2005)	N	Y	Y	Y
Southcott et al., (2010)	Y	Y	Y	N
Bailey et al., (2013)	Y	Y	Y	Y
Adeli et al., (2015)	Y	Y	Y	Y
Rieger et al., (2016)	N	NA	NA	N
Karbasy et al., (2016)	Y	Y	Y	Y
Parkin et al., (2017)	Y	Y	Y	Y
Higgins et al., (2018)	Y	Y	Y	Y
Hoq et al., (2019)	Y	NA	NA	N
Larsson et al., (2019)	N	Y	Y	Y
Bohn et al., (2019)	Y	Y	Y	Y
Bogner et al., (2019)	Y	Y	N	Y
Tahmasebi et al., (2020)	Y	Y	Y	Y
Hall et al., (2021)	Y	NA	NA	Y
Bohn et al., (2021)	Y	Y	Y	Y
Wilson et al., (2023)	Y	NA	NA	N

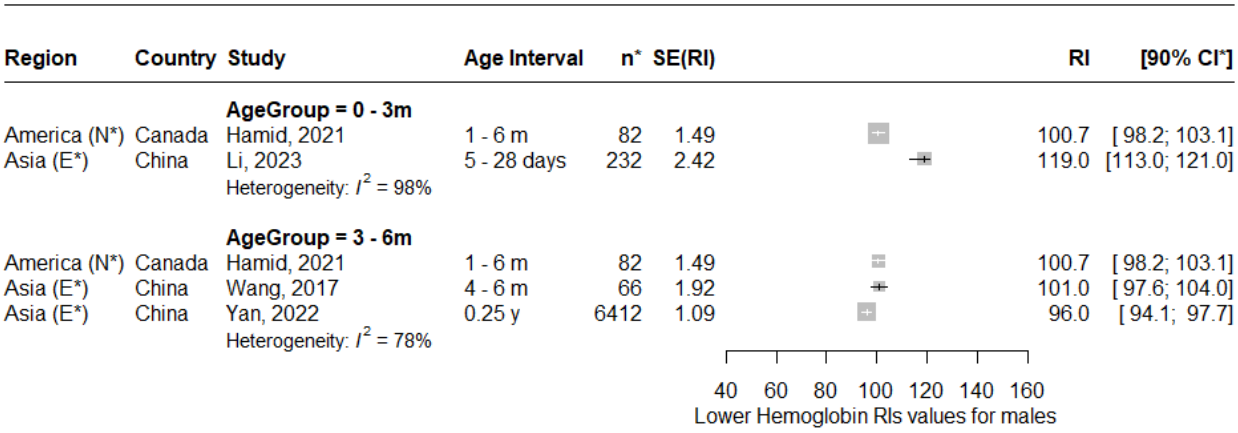


Figure S1. Forest plot of hemoglobin RIs lower limits for 0-6 months old males

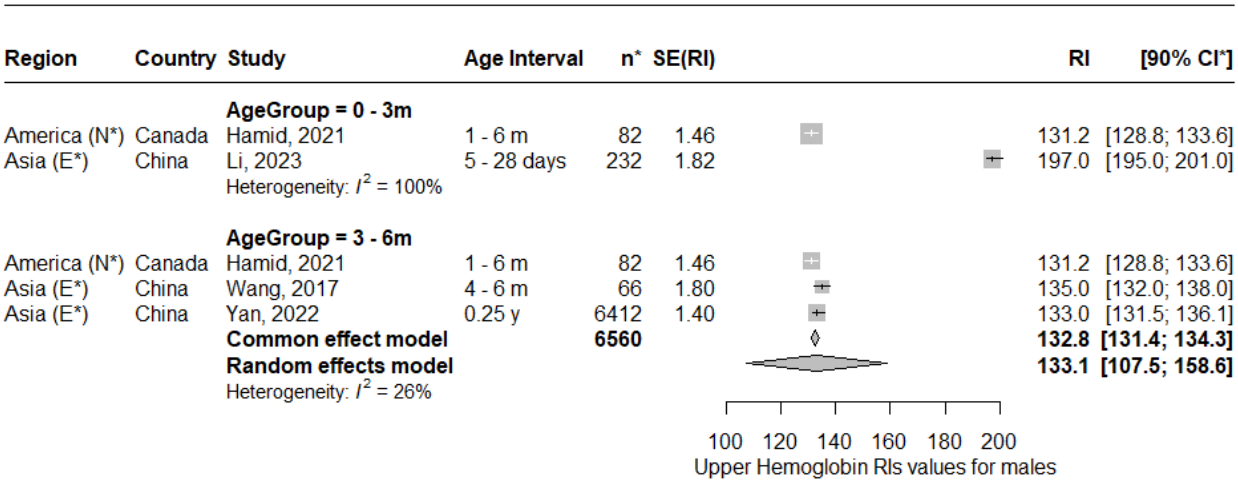


Figure S2. Forest plot of hemoglobin RIs upper limits for 0-6 months old males

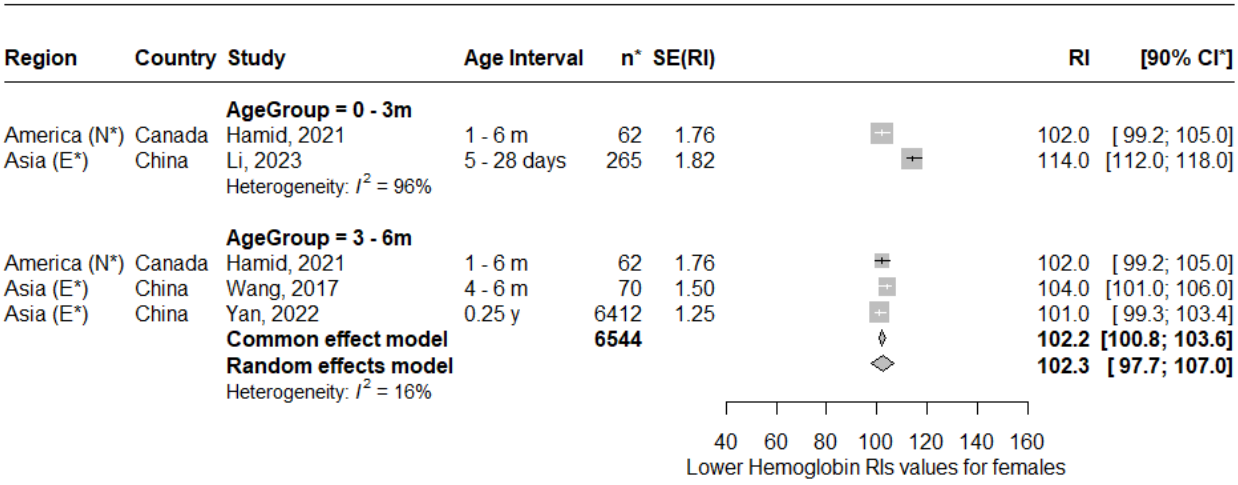


Figure S3. Forest plot of hemoglobin RIs lower limits for 0-6 months old females

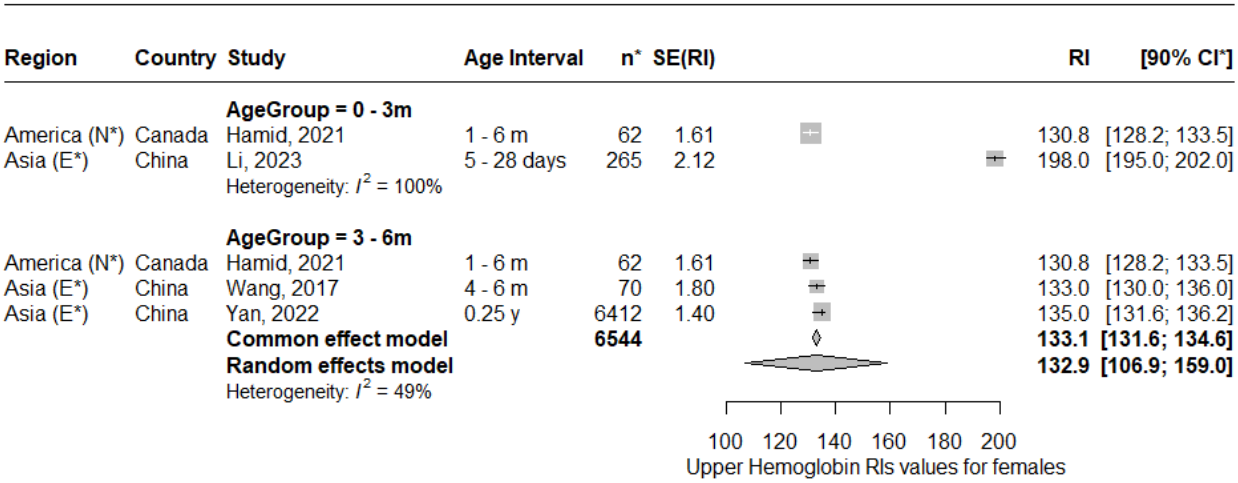


Figure S4. Forest plot of hemoglobin RIs upper limits for 0-6 months old females

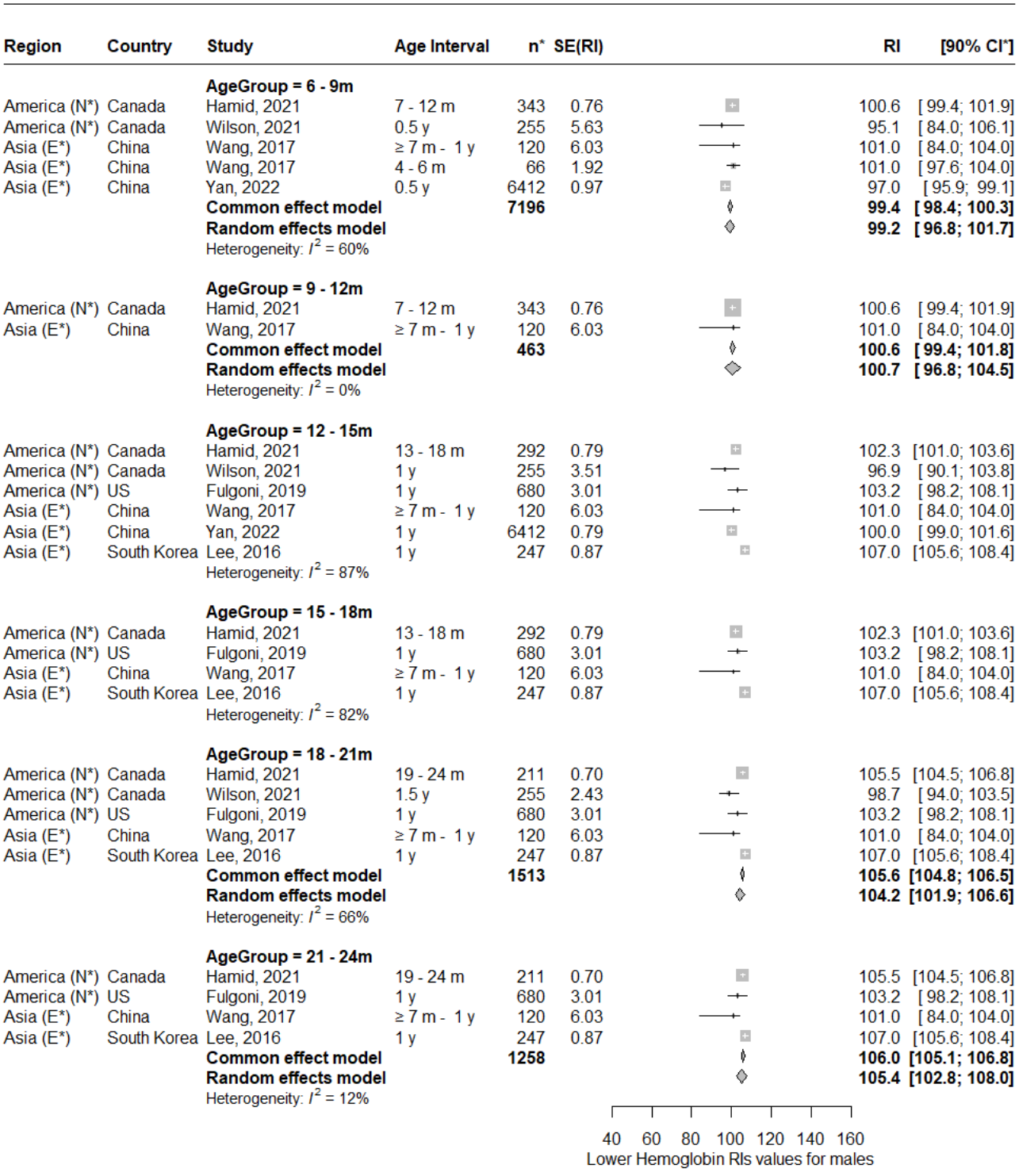


Figure S5. Forest plot of hemoglobin RIs lower limits for 6-24 months old males

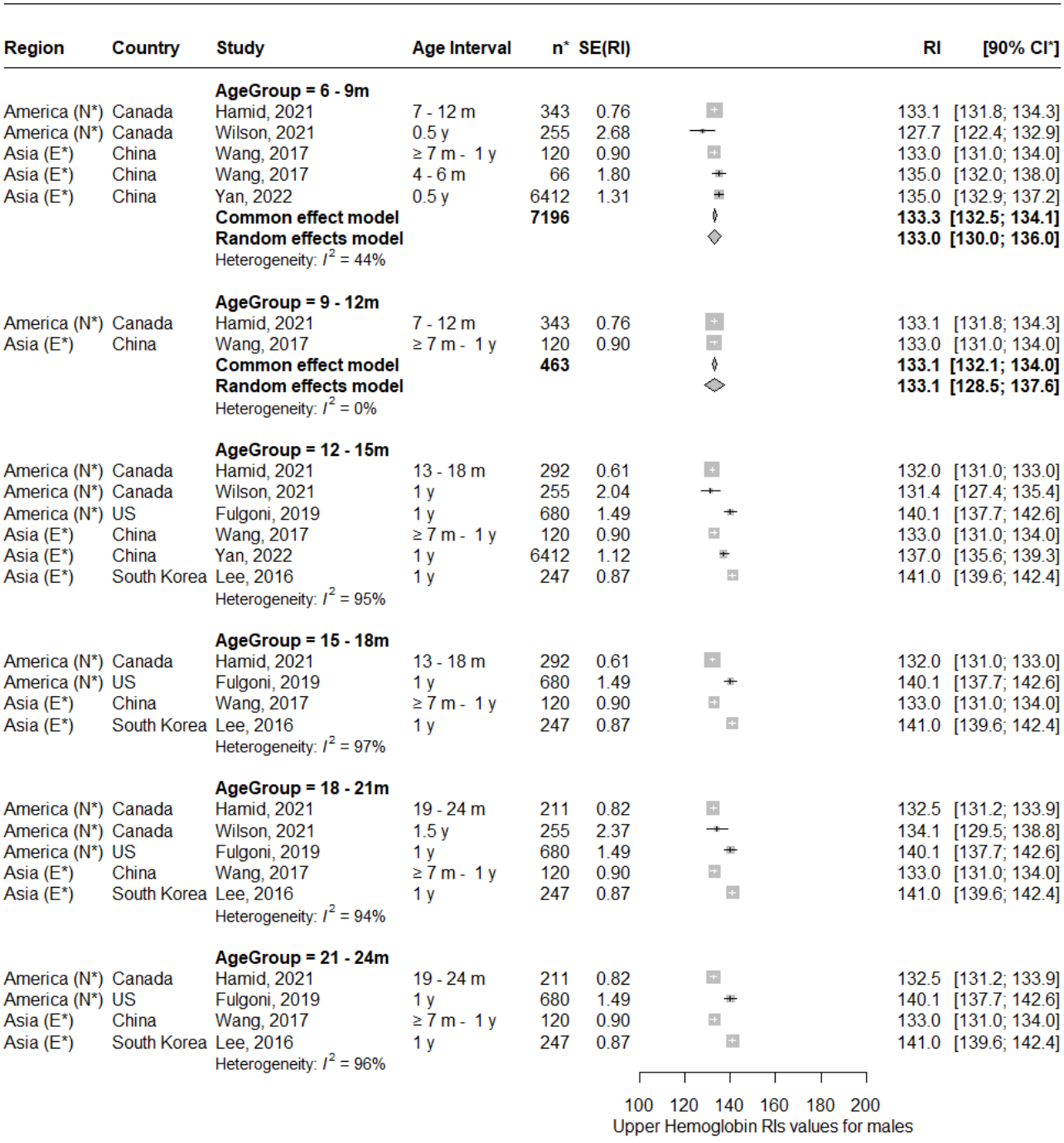


Figure S6. Forest plot of hemoglobin RIs upper limits for 6-24 months old males

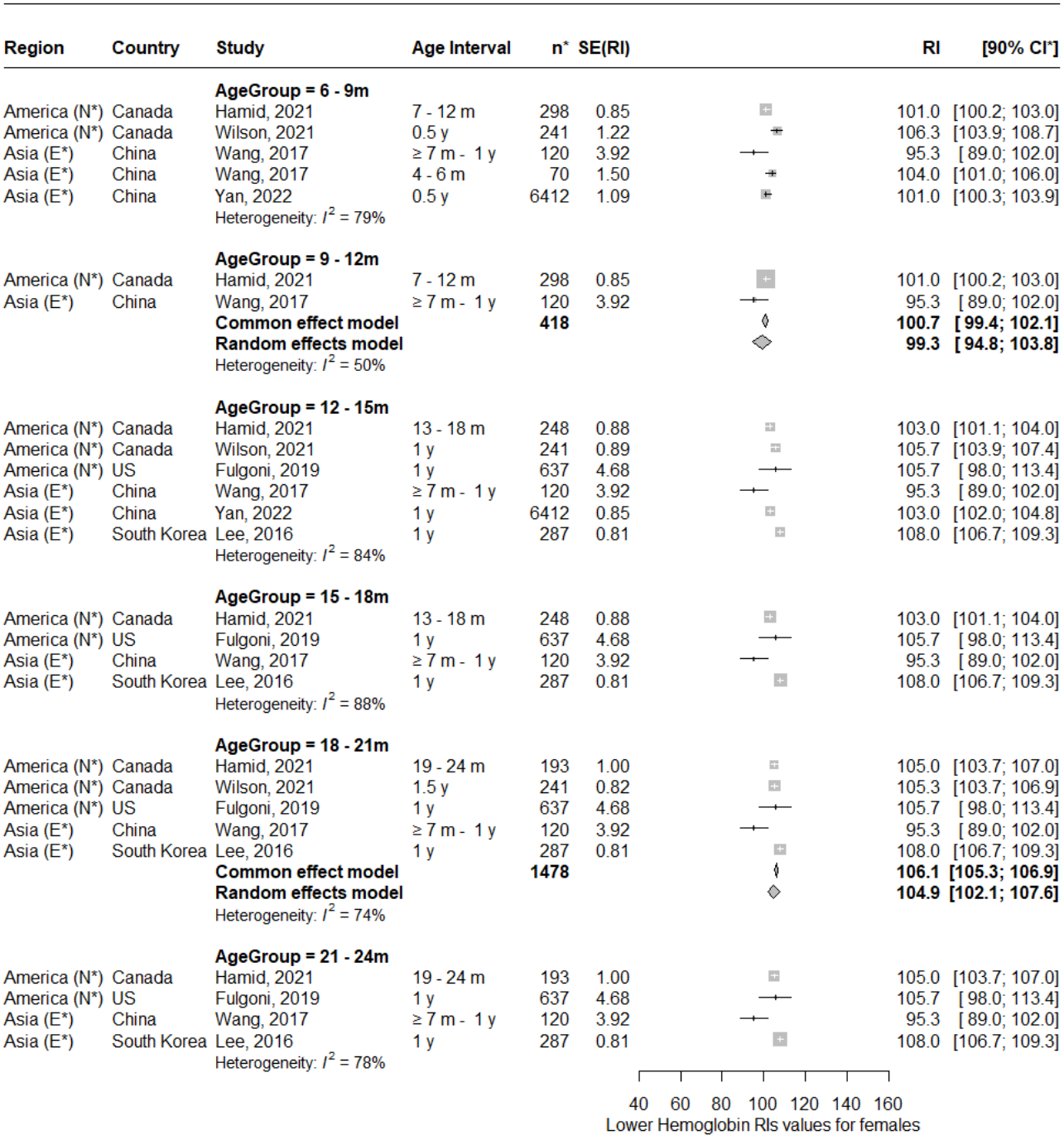


Figure S7. Forest plot of hemoglobin RIs lower limits for 6-24 months old females

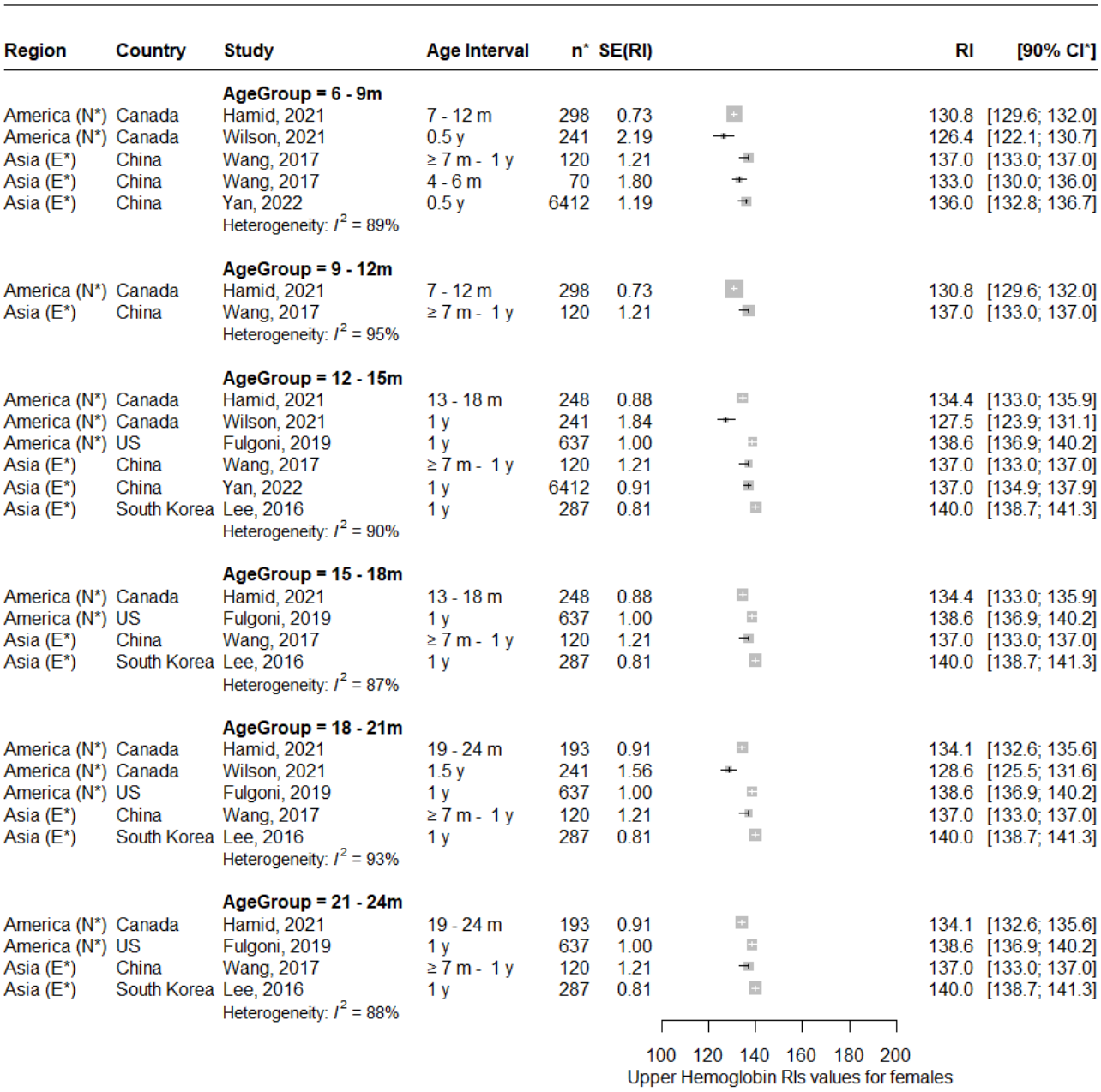


Figure S8. Forest plot of hemoglobin RIs upper limits for 6-24 months old females

Region	Country	Study	Age Interval	n*	SE(RI)	RI	[90% CI*]
		AgeGroup = 24 - 27m					
America (N*)	Canada	Hamid, 2021	25 - 30 m	113	1.03	107.3	[105.6; 109.0]
America (N*)	Canada	Wilson, 2021	2 y	255	2.01	100.5	[96.5; 104.4]
America (N*)	US	Fulgoni, 2019	2 y	714	1.49	106.6	[104.1; 109.0]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
Asia (E*)	China	Yan, 2022	2 y	6412	0.58	105.0	[103.9; 105.8]
		Heterogeneity: I^2 = 98%					
		AgeGroup = 27 - 30m					
America (N*)	Canada	Hamid, 2021	25 - 30 m	113	1.03	107.3	[105.6; 109.0]
America (N*)	US	Fulgoni, 2019	2 y	714	1.49	106.6	[104.1; 109.0]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
		Heterogeneity: I^2 = 97%					
		AgeGroup = 30 - 33m					
America (N*)	Canada	Hamid, 2021	31 - 36 m	41	2.13	103.4	[99.9; 106.9]
America (N*)	Canada	Wilson, 2021	2.5 y	255	1.79	102.1	[98.6; 105.6]
America (N*)	US	Fulgoni, 2019	2 y	714	1.49	106.6	[104.1; 109.0]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
		Heterogeneity: I^2 = 96%					
		AgeGroup = 33 - 36m					
America (N*)	Canada	Hamid, 2021	31 - 36 m	41	2.13	103.4	[99.9; 106.9]
America (N*)	US	Fulgoni, 2019	2 y	714	1.49	106.6	[104.1; 109.0]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
		Heterogeneity: I^2 = 96%					
		AgeGroup = 3 - 4y					
America (N*)	Canada	Parkin, 2017	>3 - <6 y	979	0.61	107.0	[106.0; 108.0]
America (N*)	Canada	Wilson, 2021	3.5 y	255	1.59	105.1	[102.0; 108.2]
America (N*)	US	Fulgoni, 2019	3 y	531	1.37	111.7	[109.4; 113.9]
Asia (E*)	China	Sun, 2020	>3 - 5 y	2082	0.61	118.0	[117.0; 119.0]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
Asia (E*)	China	Yan, 2022	3 y	6412	0.61	108.0	[107.2; 109.2]
		Heterogeneity: I^2 = 98%					
		AgeGroup = 4 - 5y					
America (N*)	Canada	Parkin, 2017	>3 - <6 y	979	0.61	107.0	[106.0; 108.0]
America (N*)	Canada	Wilson, 2021	4.5 y	255	1.71	107.8	[104.5; 111.2]
America (N*)	US	Fulgoni, 2019	4 y	552	1.52	112.0	[109.5; 114.5]
Asia (E*)	China	Sun, 2020	>3 - 5 y	2082	0.61	118.0	[117.0; 119.0]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
Asia (E*)	China	Yan, 2022	4 y	6412	0.64	111.0	[109.5; 111.6]
		Heterogeneity: I^2 = 97%					

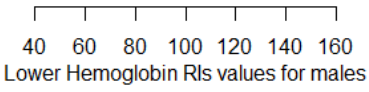


Figure S9. Forest plot of hemoglobin RIs lower limits for 2-5 years old males

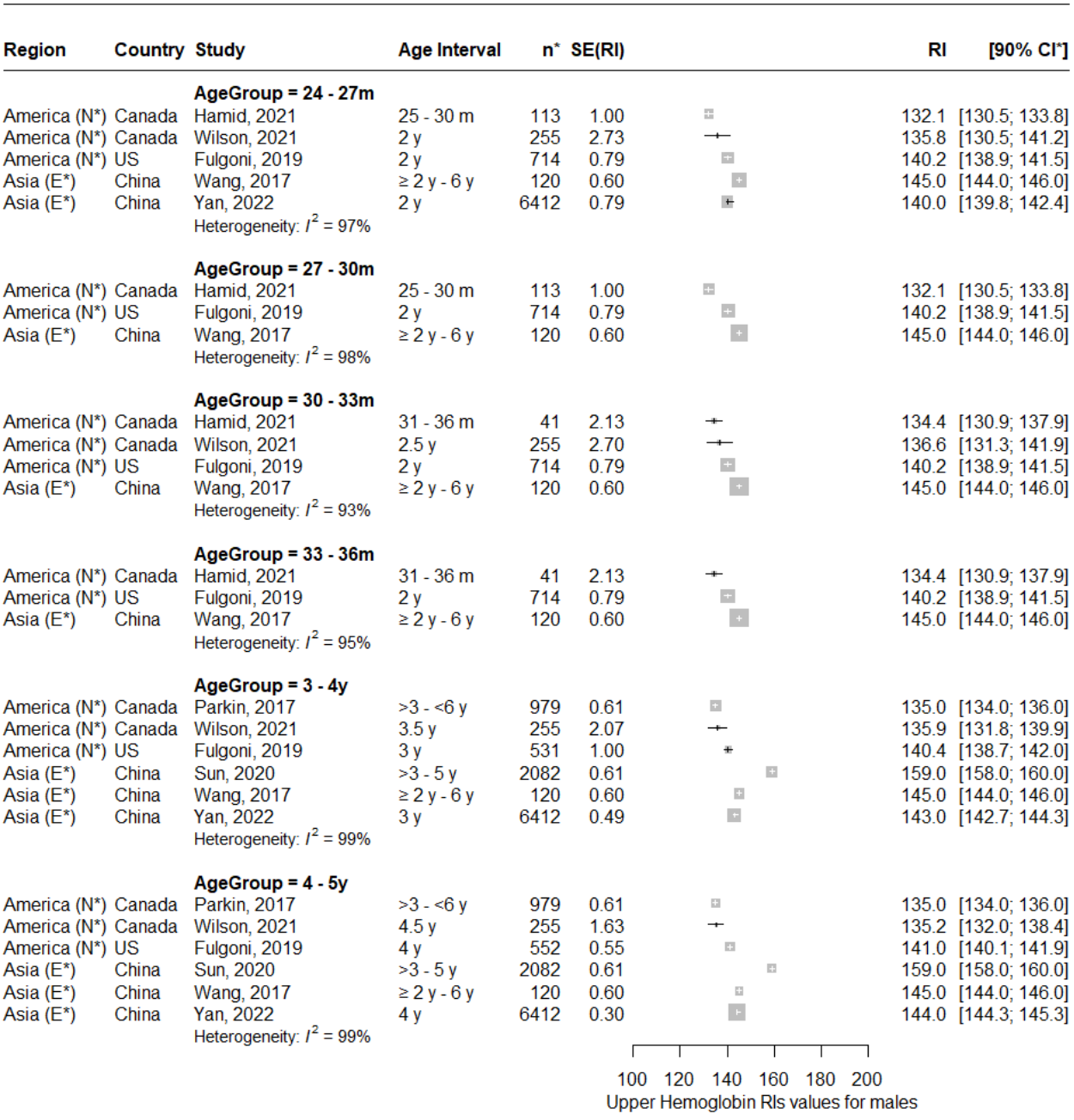


Figure S10. Forest plot of hemoglobin RIs upper limits for 2-5 years old males

Region	Country	Study	Age Interval	n*	SE(RI)	RI	[90% CI*]
		AgeGroup = 24 - 27m					
America (N*)	Canada	Hamid, 2021	25 - 30 m	101	1.31	105.0	[102.7; 107.0]
America (N*)	Canada	Wilson, 2021	2 y	241	0.92	105.0	[103.2; 106.8]
America (N*)	US	Fulgoni, 2019	2 y	716	2.13	107.3	[103.8; 110.8]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
Asia (E*)	China	Yan, 2022	2 y	6412	0.64	106.0	[105.1; 107.2]
		Heterogeneity: I^2 = 97%					
		AgeGroup = 27 - 30m					
America (N*)	Canada	Hamid, 2021	25 - 30 m	101	1.31	105.0	[102.7; 107.0]
America (N*)	US	Fulgoni, 2019	2 y	716	2.13	107.3	[103.8; 110.8]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
		Heterogeneity: I^2 = 96%					
		AgeGroup = 30 - 33m					
America (N*)	Canada	Hamid, 2021	31 - 36 m	28	1.82	108.0	[105.0; 111.0]
America (N*)	Canada	Wilson, 2021	2.5 y	241	1.05	104.9	[102.8; 106.9]
America (N*)	US	Fulgoni, 2019	2 y	716	2.13	107.3	[103.8; 110.8]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
		Heterogeneity: I^2 = 96%					
		AgeGroup = 33 - 36m					
America (N*)	Canada	Hamid, 2021	31 - 36 m	28	1.82	108.0	[105.0; 111.0]
America (N*)	US	Fulgoni, 2019	2 y	716	2.13	107.3	[103.8; 110.8]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
		Heterogeneity: I^2 = 91%					
		AgeGroup = 3 - 4y					
America (N*)	Canada	Parkin, 2017	>3 - <6 y	844	0.61	108.0	[107.0; 109.0]
America (N*)	Canada	Wilson, 2021	3.5 y	241	1.33	105.1	[102.5; 107.7]
America (N*)	US	Fulgoni, 2019	3 y	444	2.40	107.7	[103.8; 111.7]
Asia (E*)	China	Sun, 2020	>3 - 5 y	1952	0.61	118.0	[117.0; 119.0]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
Asia (E*)	China	Yan, 2022	3 y	6412	0.64	108.0	[107.6; 109.7]
		Heterogeneity: I^2 = 98%					
		AgeGroup = 4 - 5y					
America (N*)	Canada	Parkin, 2017	>3 - <6 y	844	0.61	108.0	[107.0; 109.0]
America (N*)	Canada	Wilson, 2021	4.5 y	241	1.48	105.9	[103.0; 108.8]
America (N*)	US	Fulgoni, 2019	4 y	511	1.09	112.3	[110.5; 114.1]
Asia (E*)	China	Sun, 2020	>3 - 5 y	1952	0.61	118.0	[117.0; 119.0]
Asia (E*)	China	Wang, 2017	≥ 2 y - 6 y	120	0.60	115.0	[114.0; 116.0]
Asia (E*)	China	Yan, 2022	4 y	6412	0.67	110.0	[109.5; 111.7]
		Heterogeneity: I^2 = 97%					

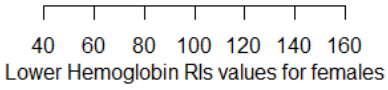


Figure S11. Forest plot of hemoglobin RIs lower limits for 2-5 years old females

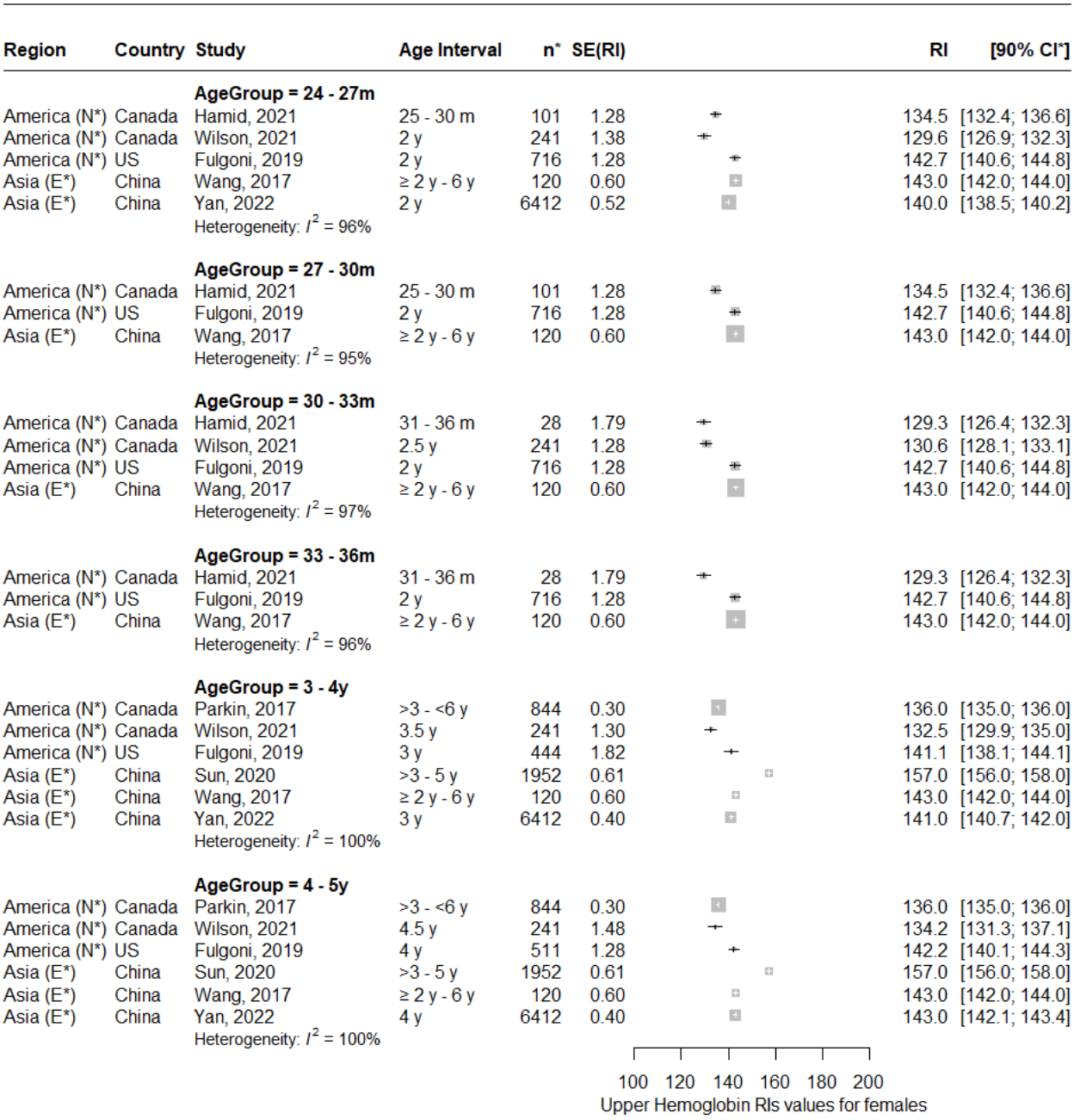


Figure S12. Forest plot of hemoglobin RIs upper limits for 2-5 years old females

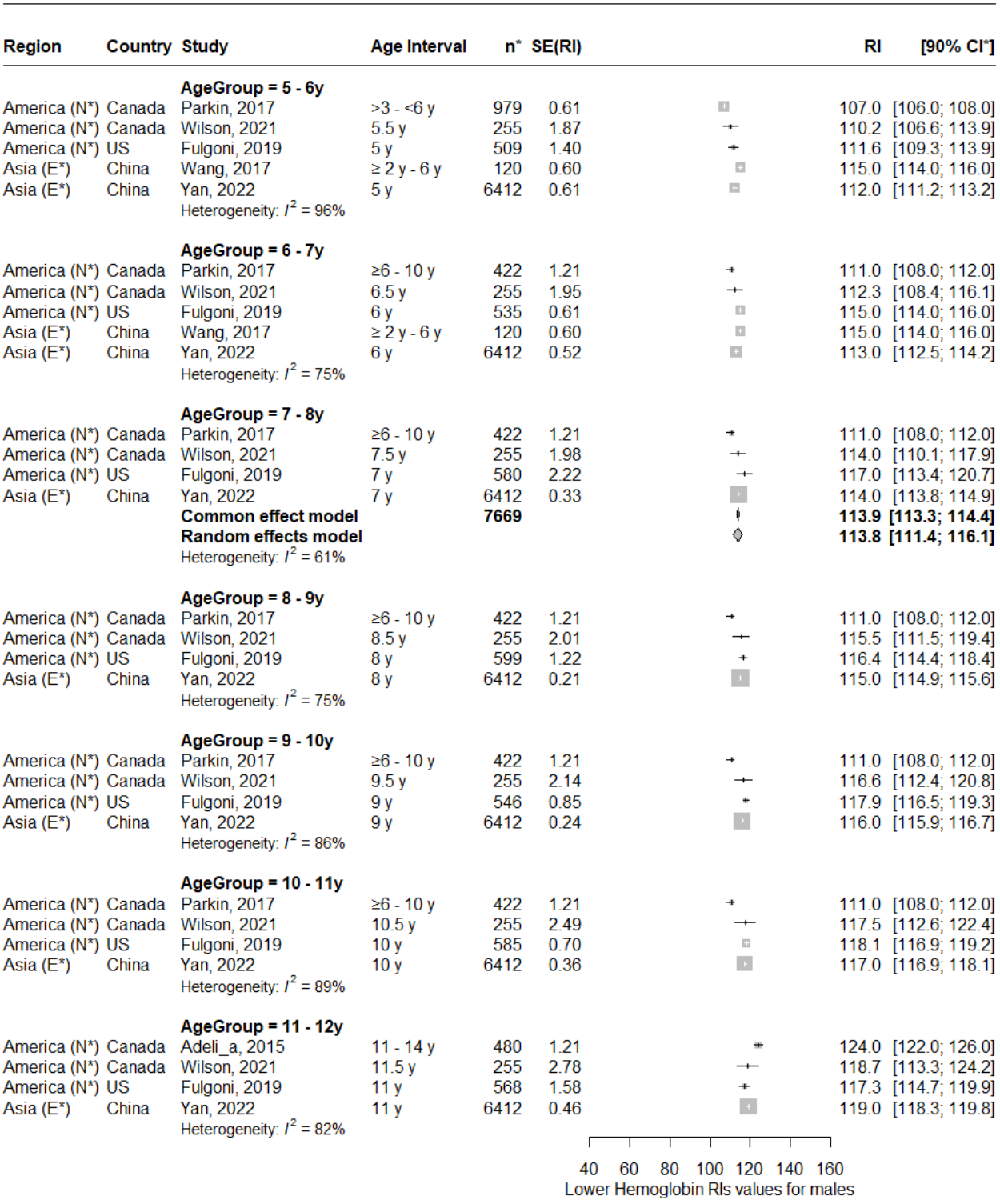


Figure S13. Forest plot of hemoglobin RIs lower limits for 5-12 years old males

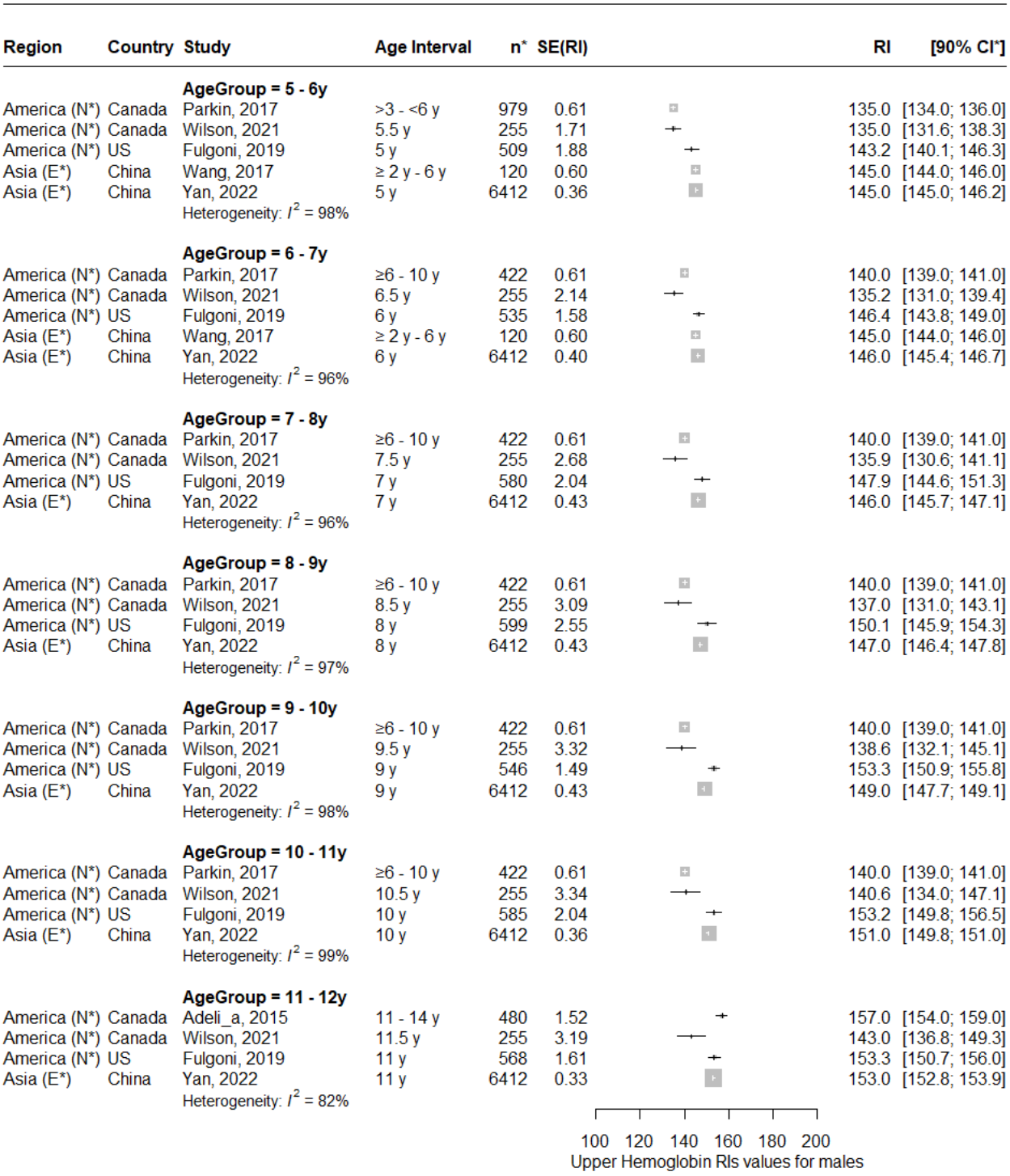


Figure S14. Forest plot of hemoglobin RIs upper limits for 5-12 years old males

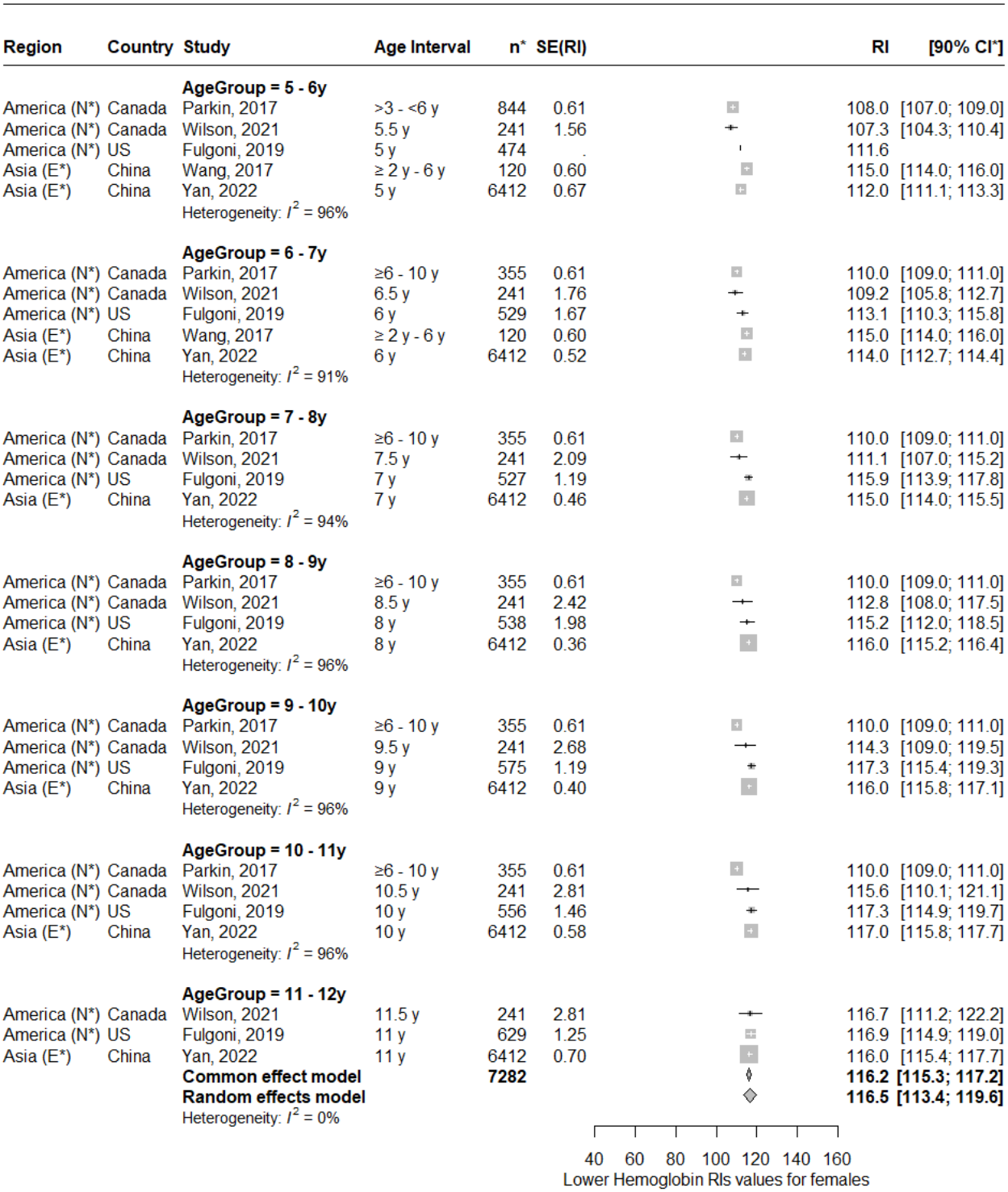


Figure S15. Forest plot of hemoglobin RIs lower limits for 5-12 years old females

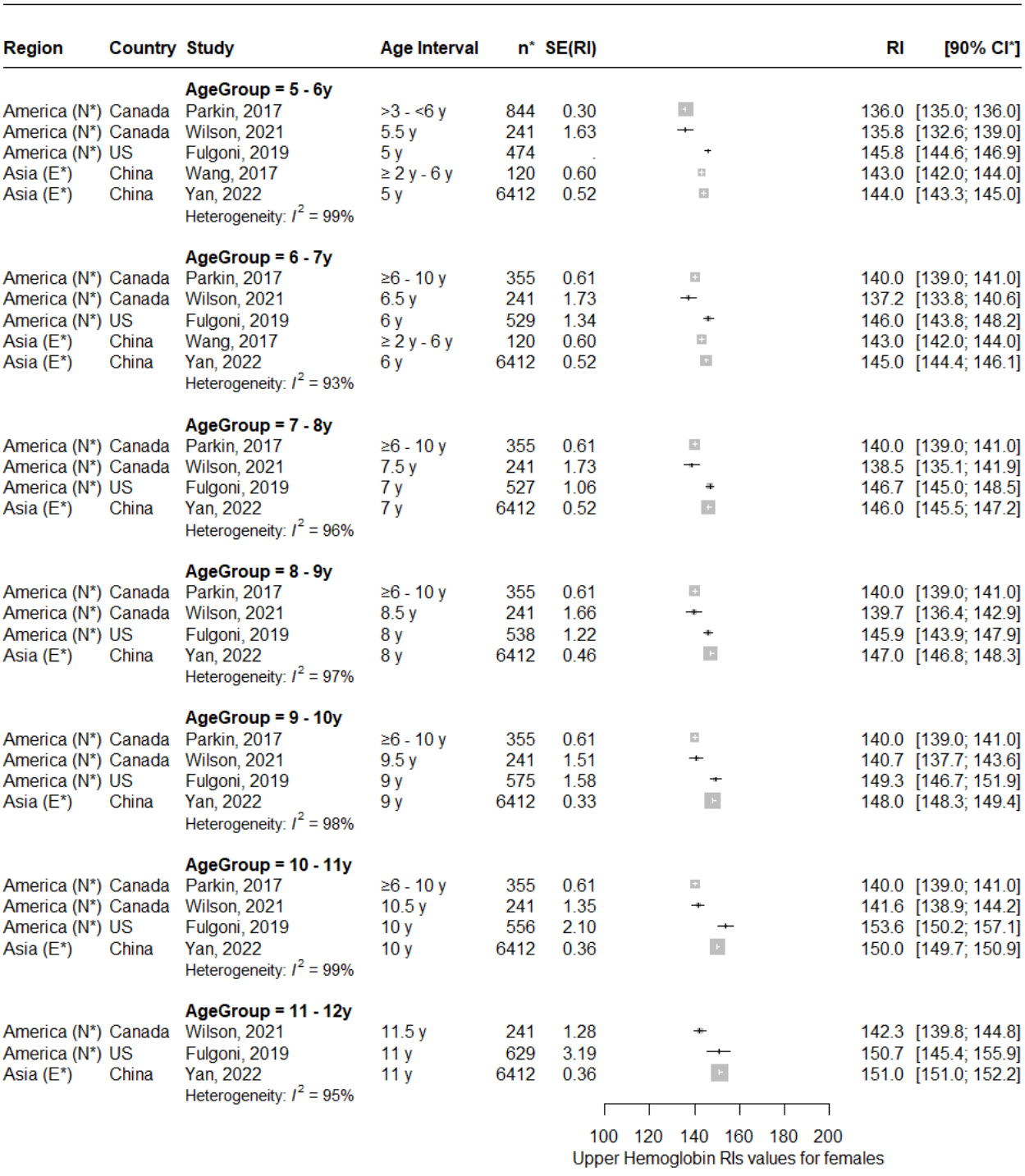


Figure S16. Forest plot of hemoglobin RIs upper limits for 5-12 years old females

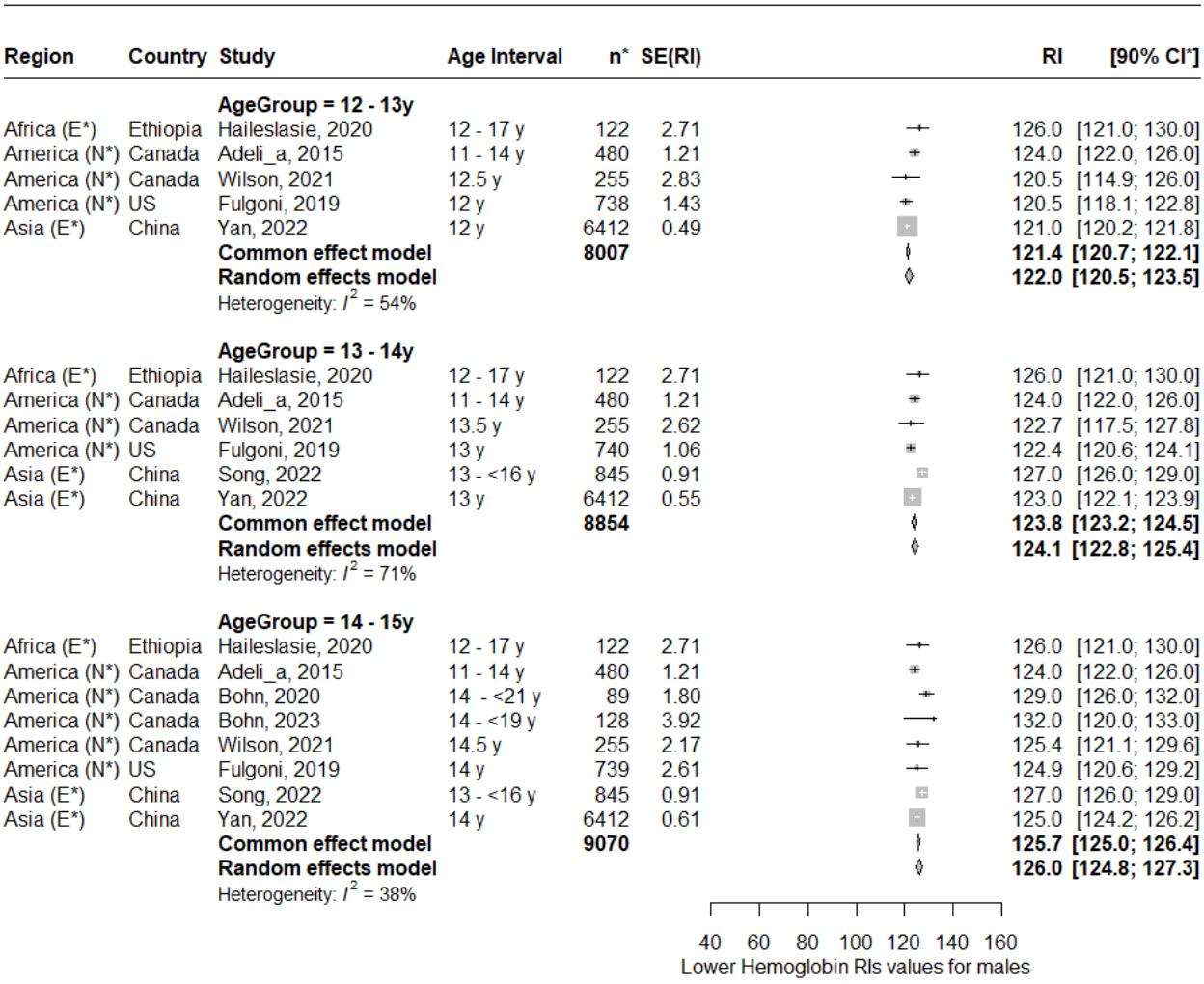


Figure S17. Forest plot of hemoglobin RIs lower limits for 12-15 years old males

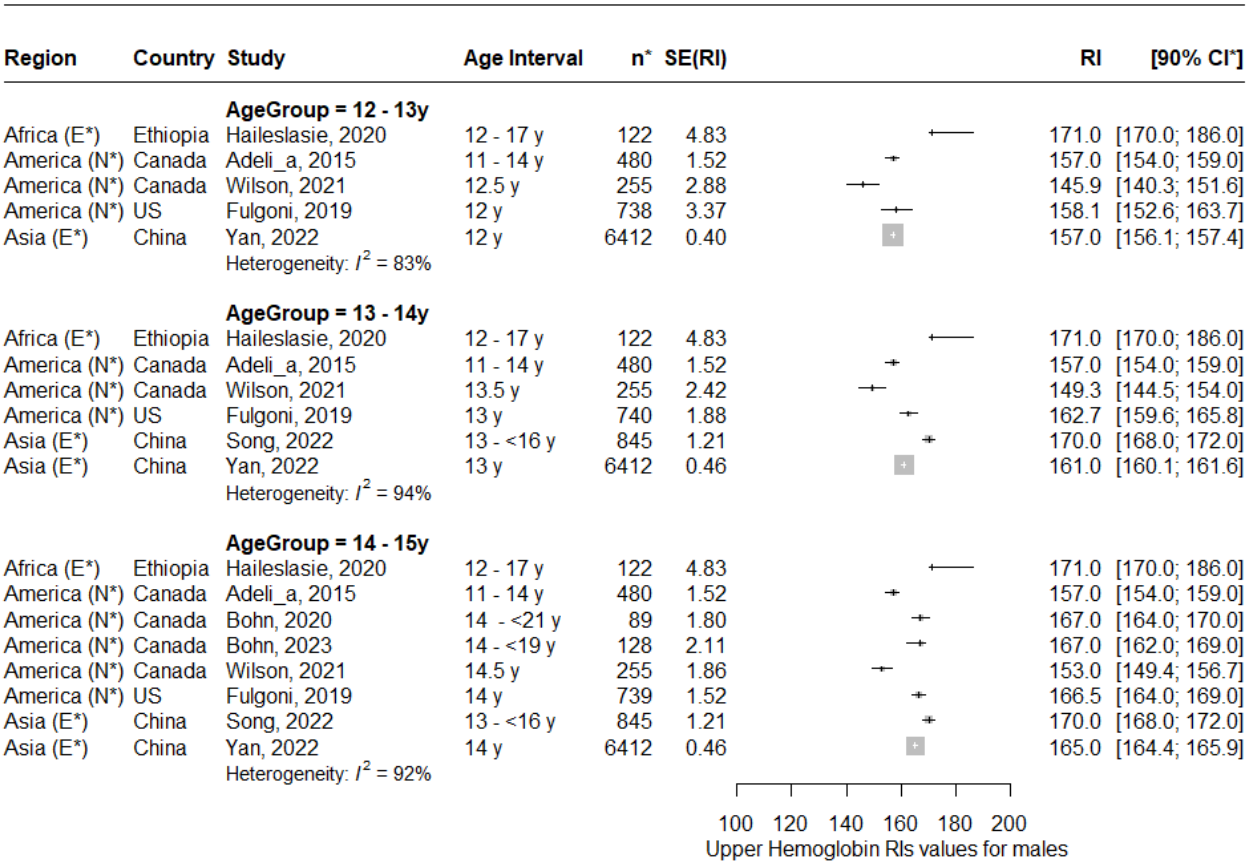


Figure S18. Forest plot of hemoglobin RIs upper limits for 12-15 years old males

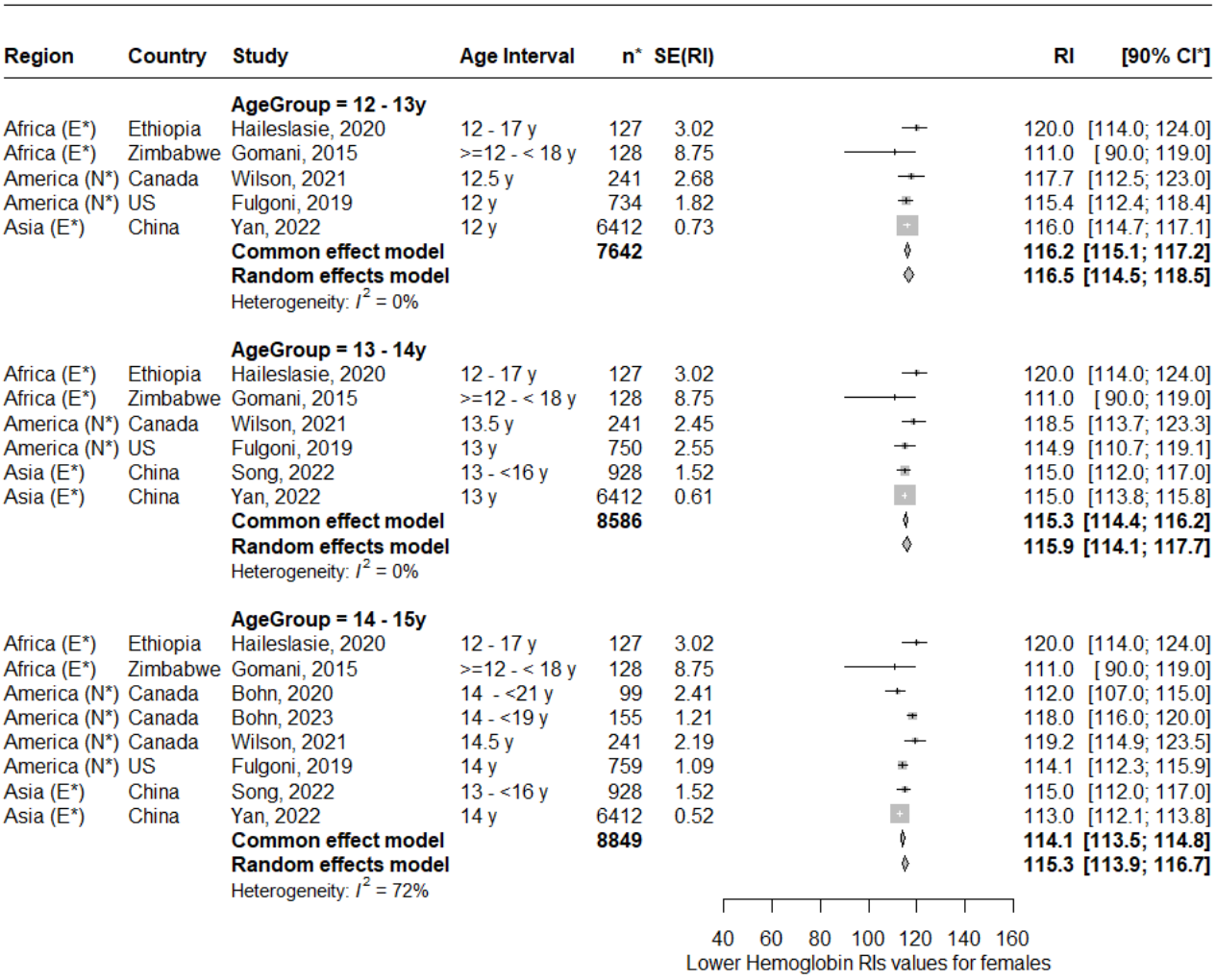


Figure S19. Forest plot of hemoglobin RIs lower limits for 12-15 years old females

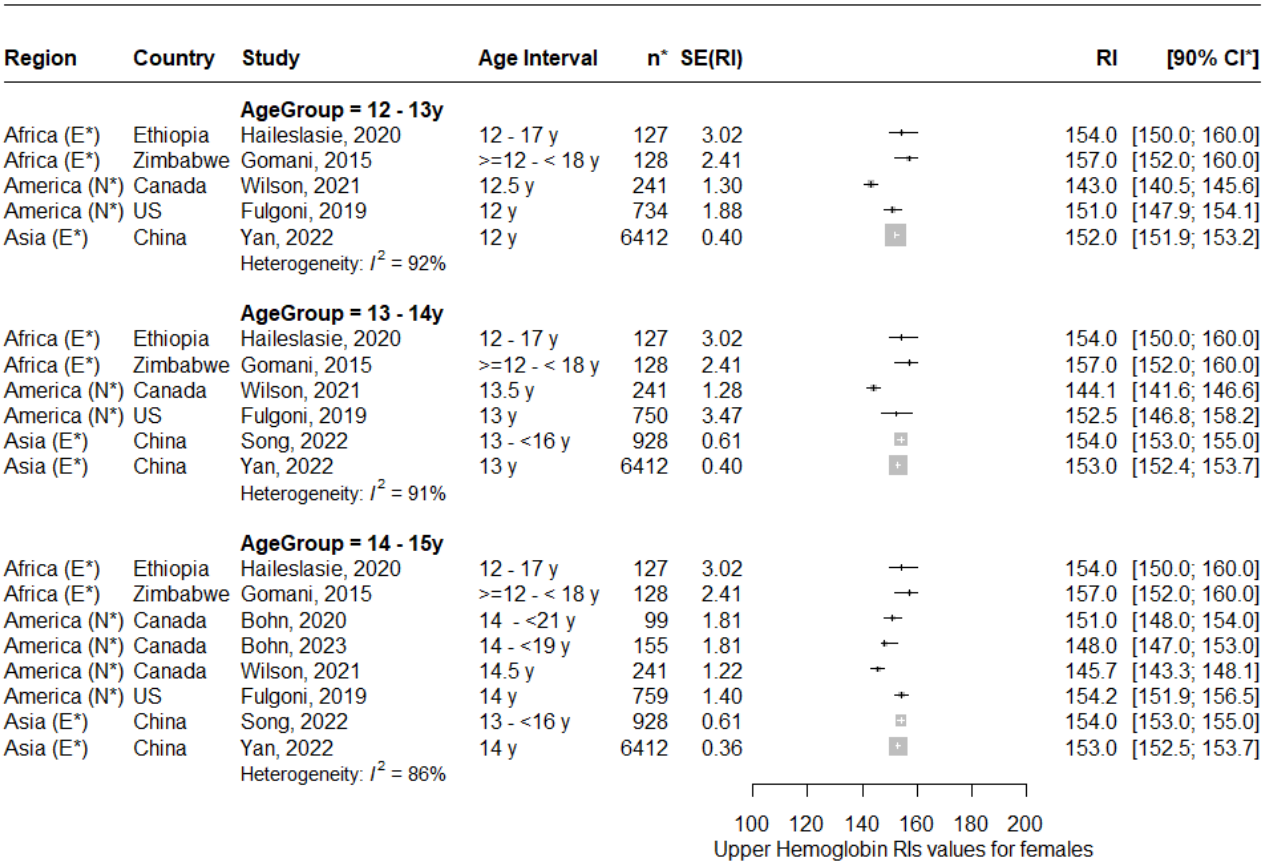


Figure S20. Forest plot of hemoglobin RIs upper limits for 12-15 years old females

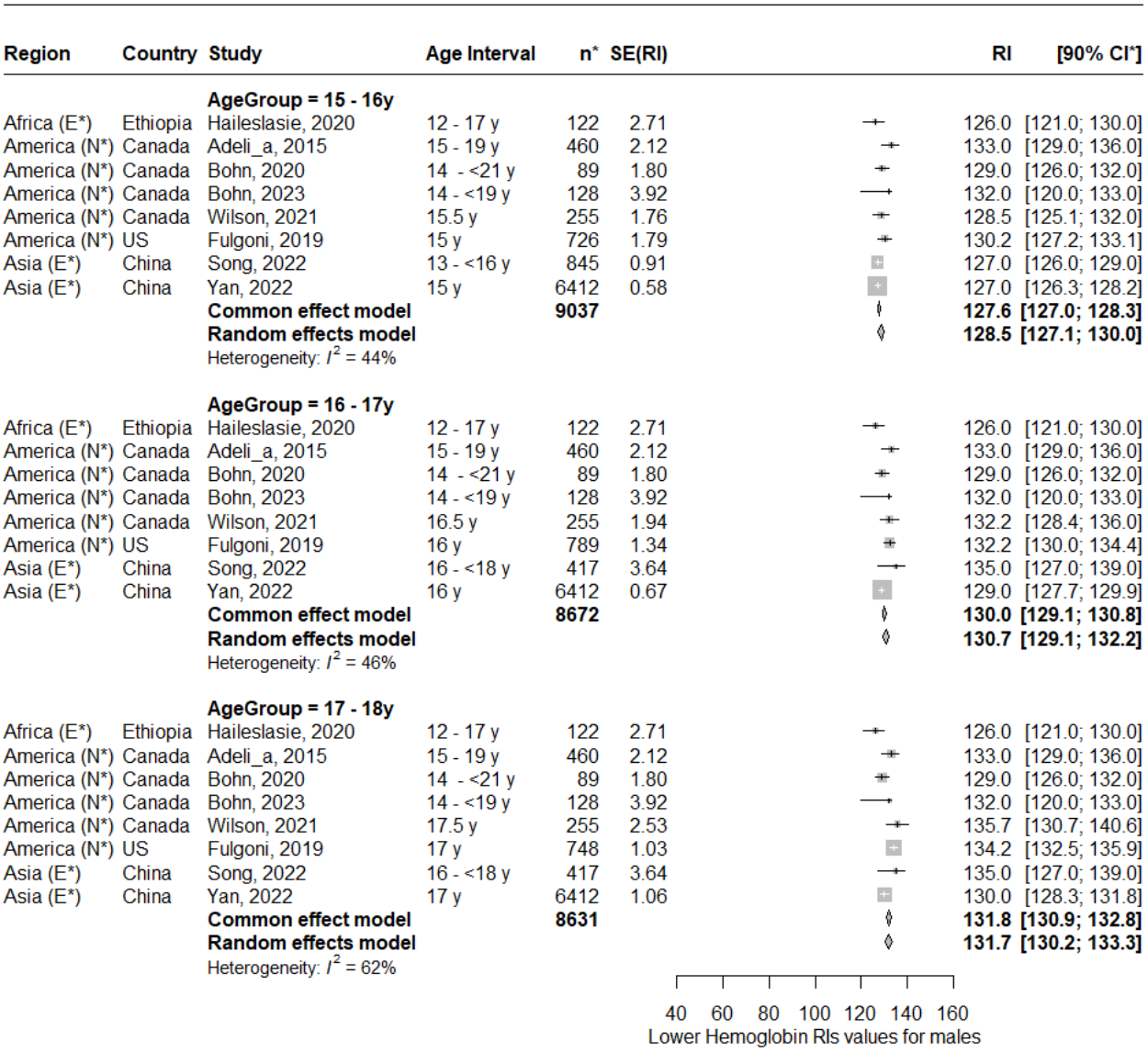


Figure S21. Forest plot of hemoglobin RIs lower limits for 15-18 years old males

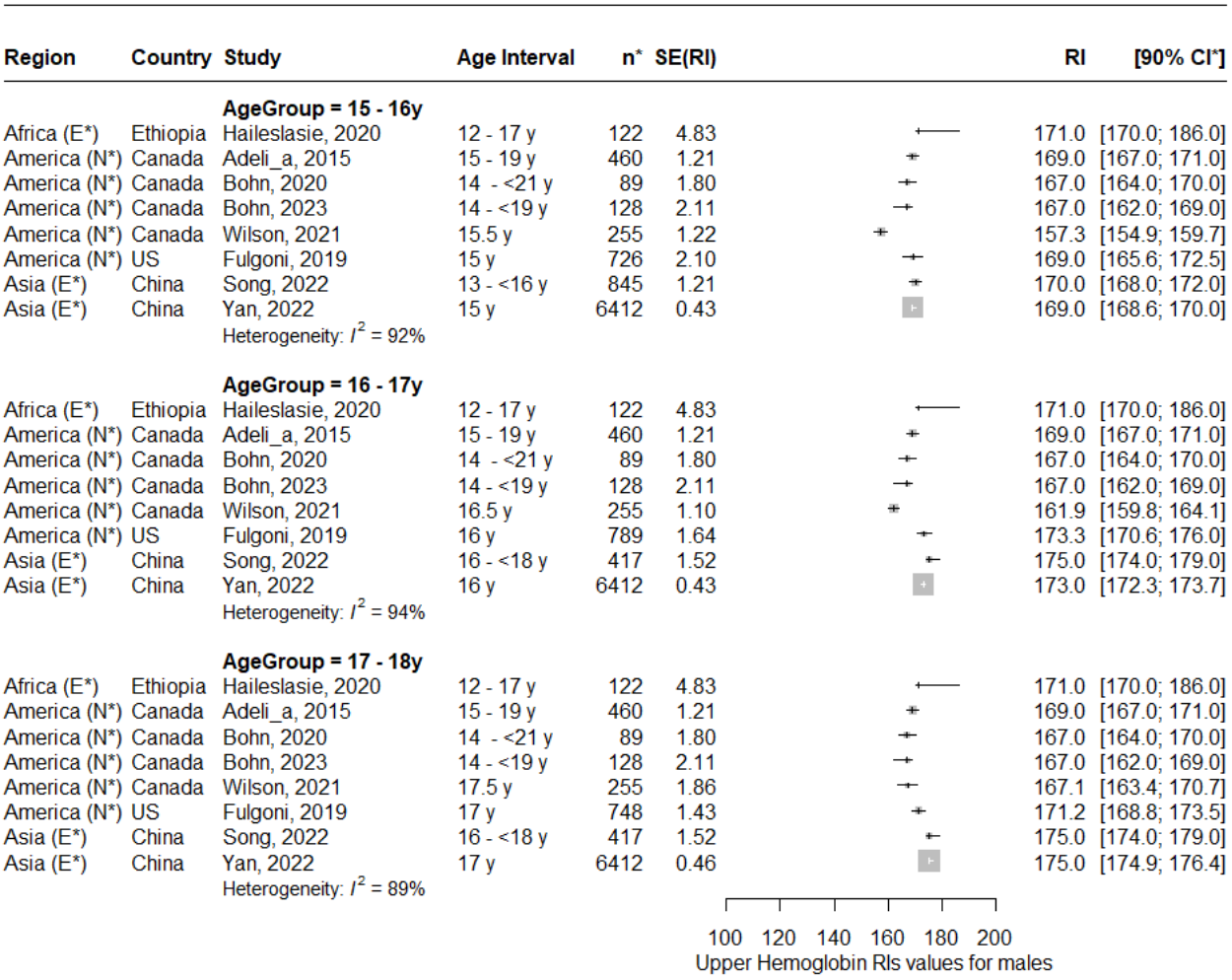


Figure S22. Forest plot of hemoglobin RIs upper limits for 15-18 years old males

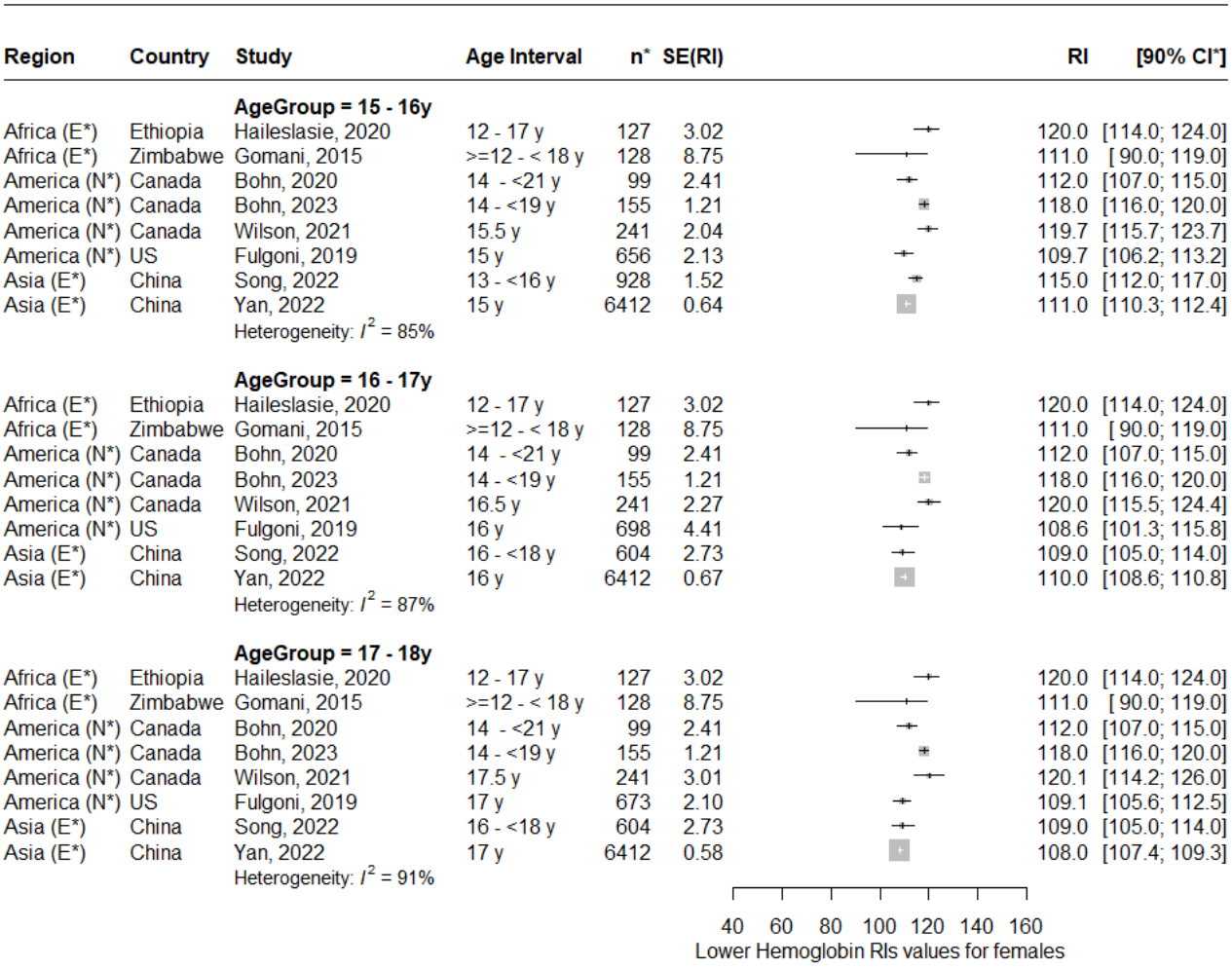


Figure S23. Forest plot of hemoglobin RIs lower limits for 15-18 years old females

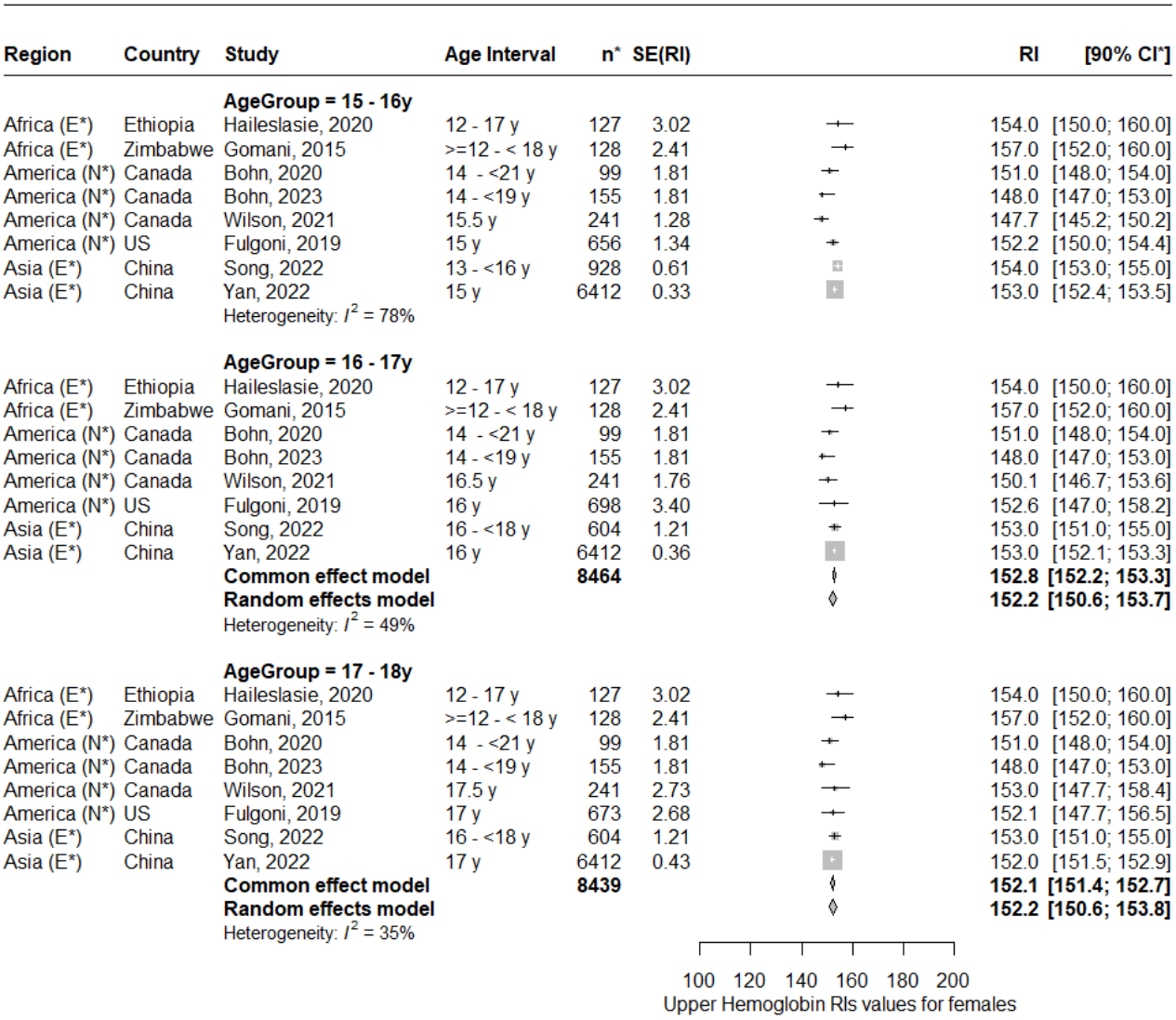


Figure S24. Forest plot of hemoglobin RIs upper limits for 15-18 years old females

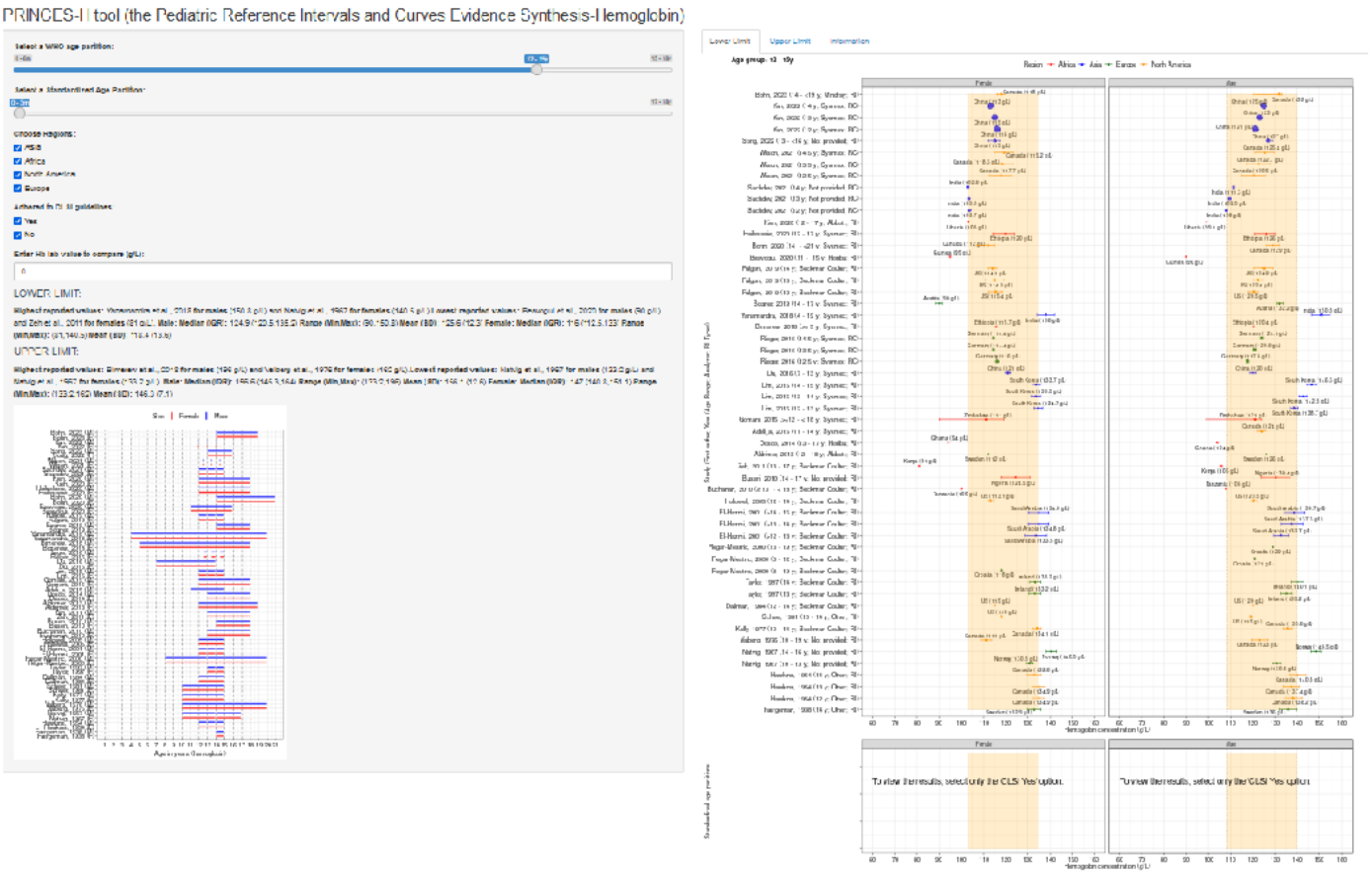


Figure S25. Screenshot of PRINCES-H tool

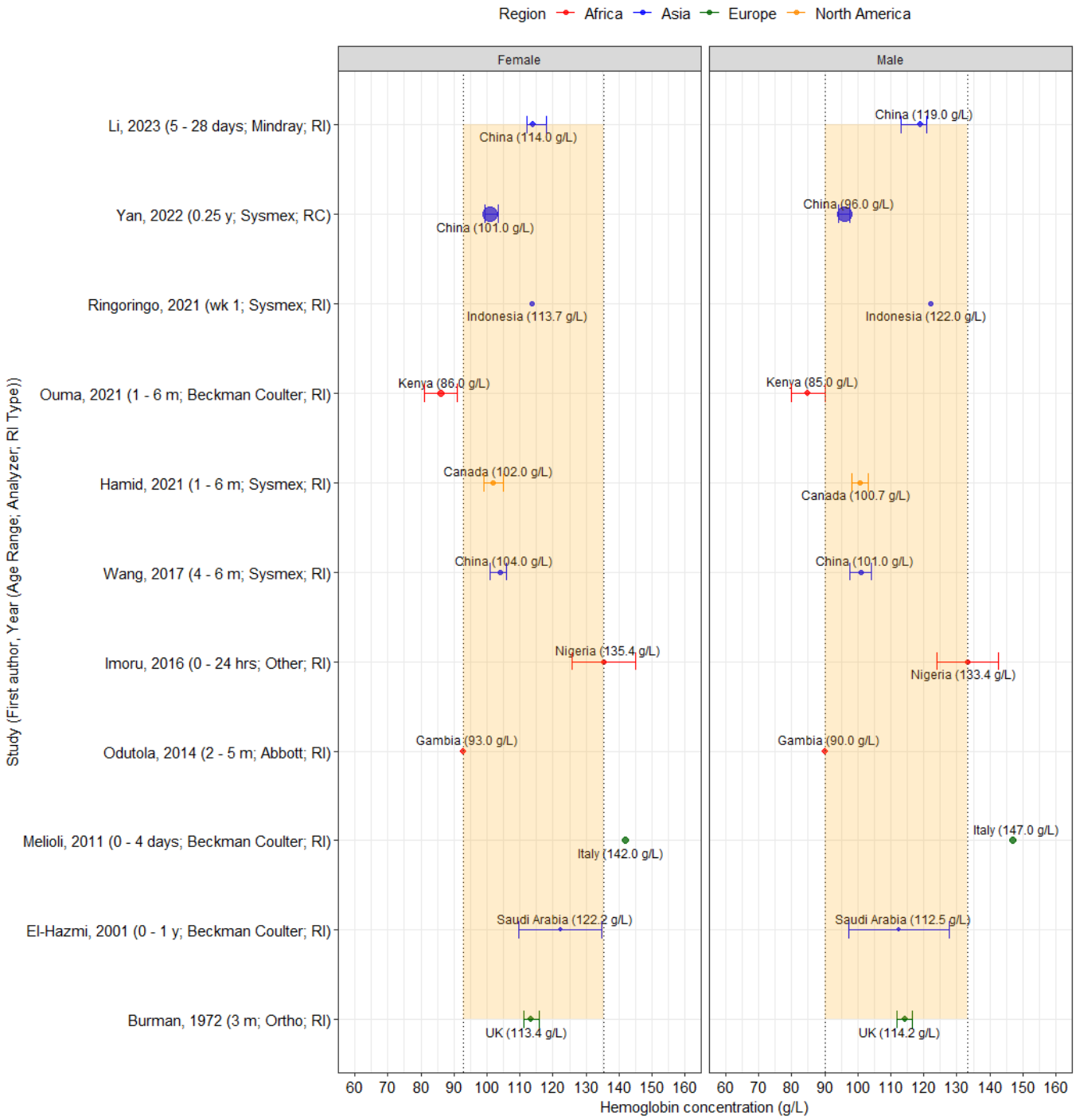


Figure S26. Distribution of lower Hemoglobin RIs limits for 0-6 months old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

Figure S27. Distribution of upper Hemoglobin RIs limits for 0-6 months old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

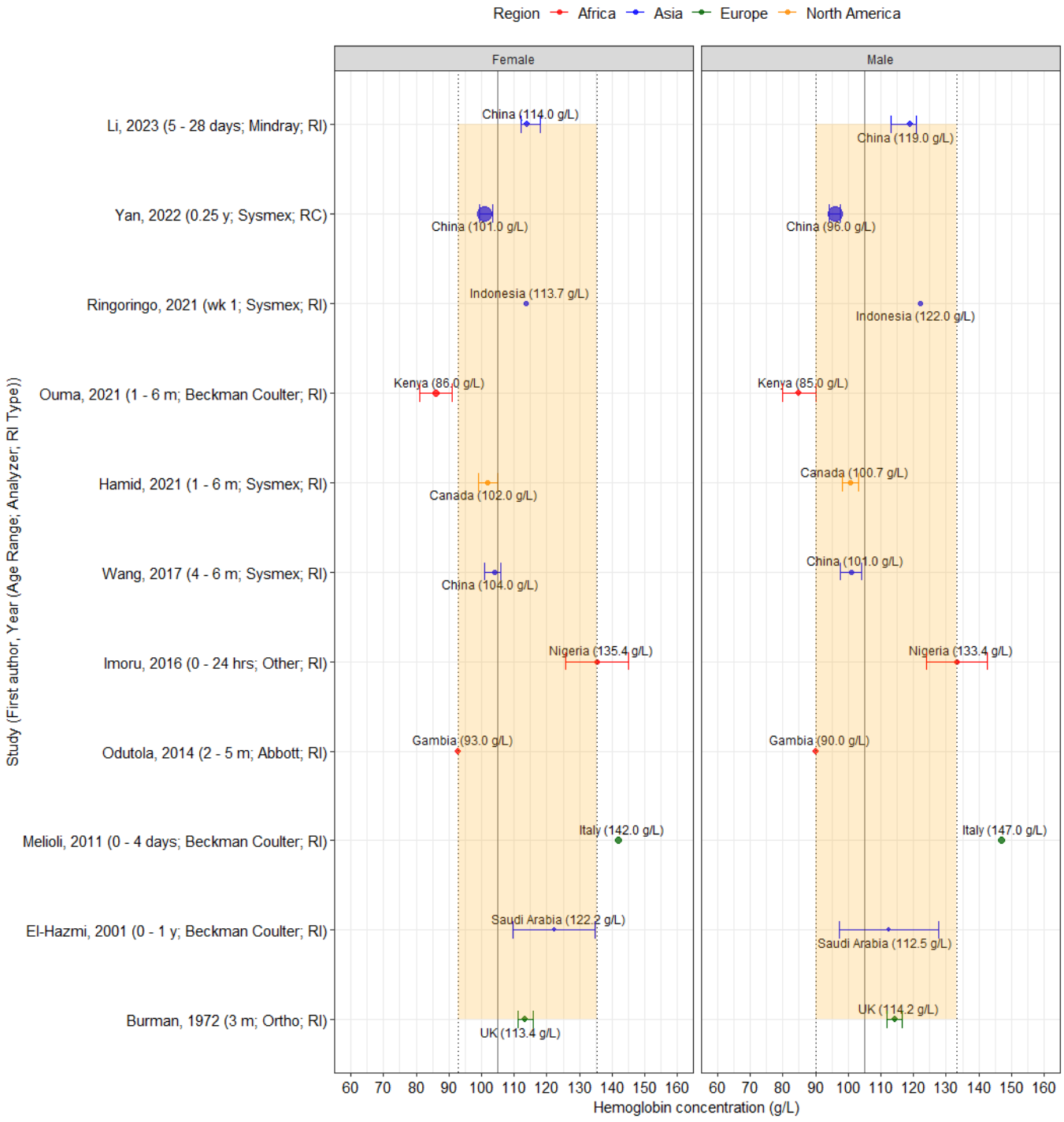


Figure S28. Distribution of lower Hemoglobin RIs limits for 6-23 months old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

Figure S29. Distribution of upper Hemoglobin RIs limits for 6-23 months old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

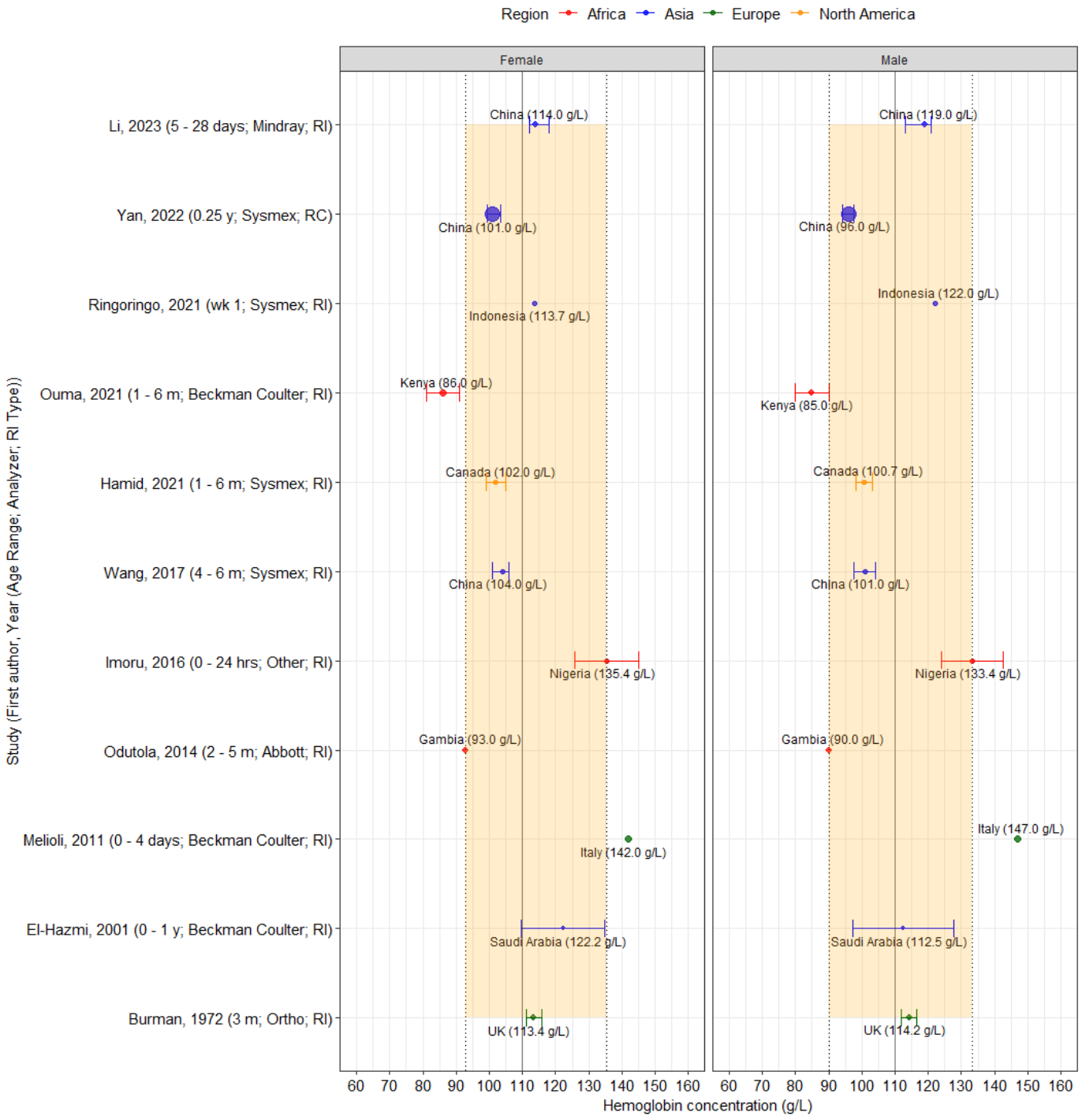


Figure S30. Distribution of lower Hemoglobin RIs limits for 24-59 months old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

Figure S31. Distribution of upper Hemoglobin RIs limits for 24-59 months old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

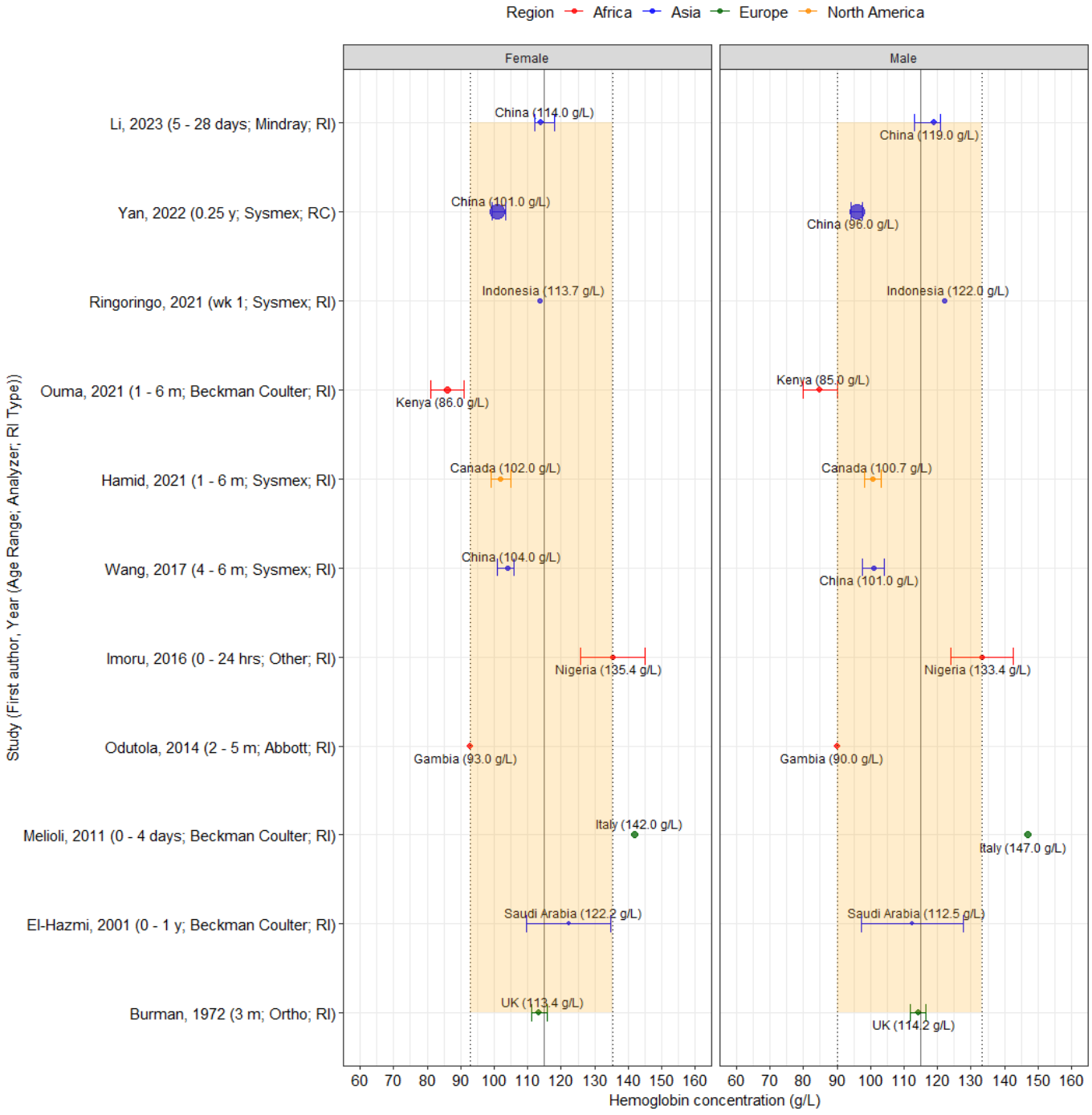


Figure S32. Distribution of lower Hemoglobin RIs limits for 5-12 years old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

Figure S33. Distribution of upper Hemoglobin RIs limits for 5-12 years old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

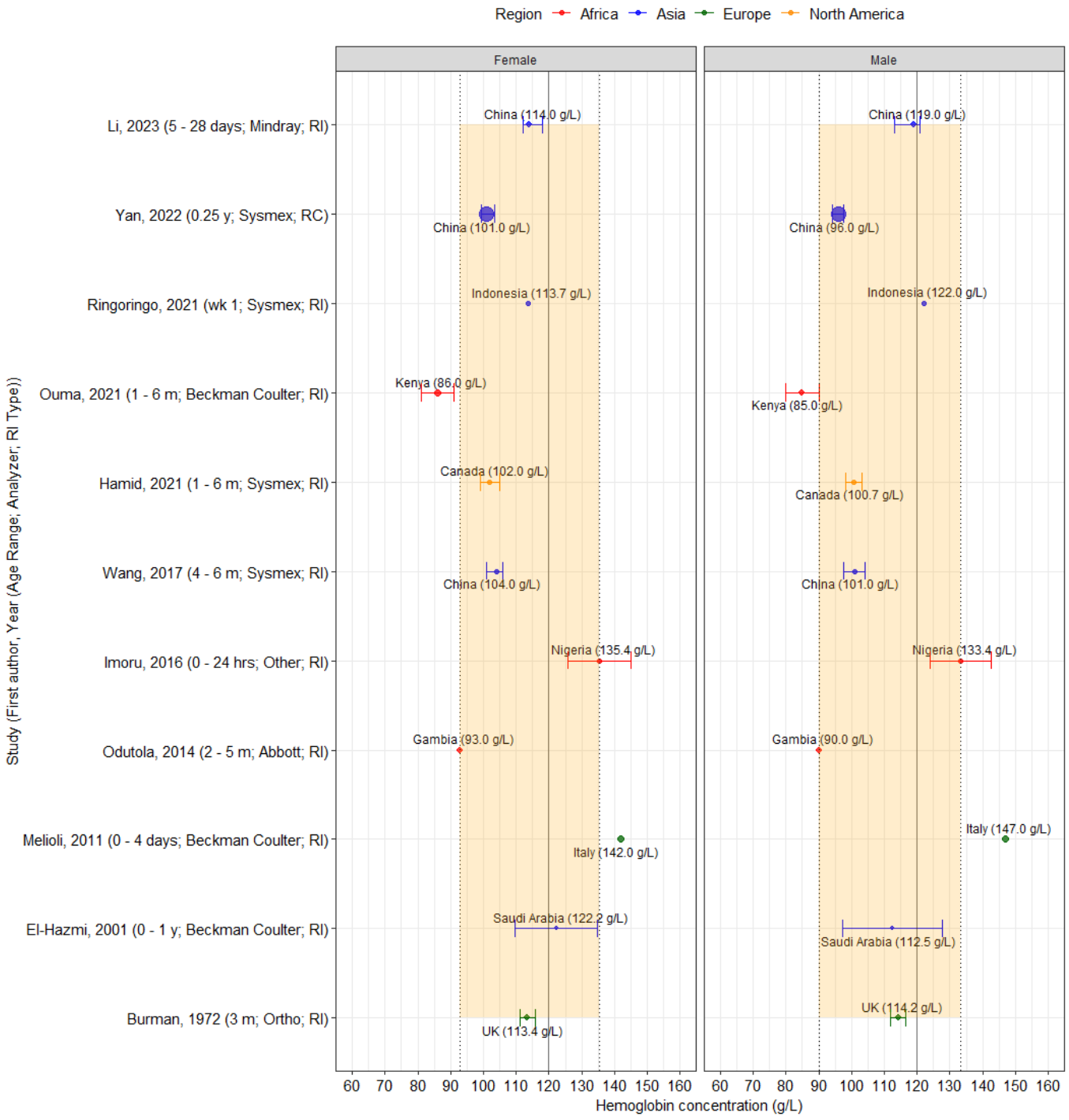


Figure S34. Distribution of lower Hemoglobin RIs limits for 12-15 years old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

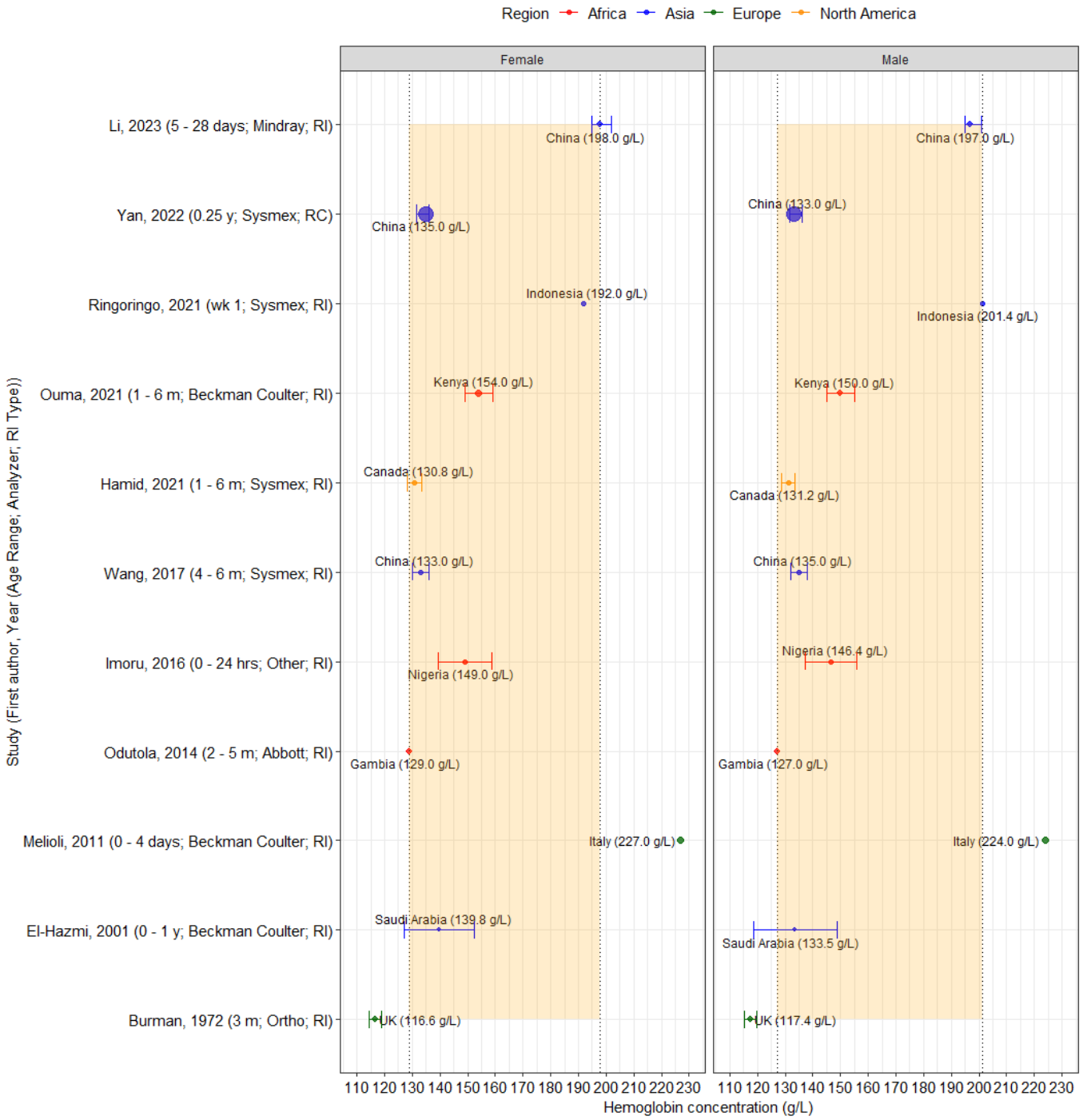


Figure S35. Distribution of upper Hemoglobin RIs limits for 12-15 years old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

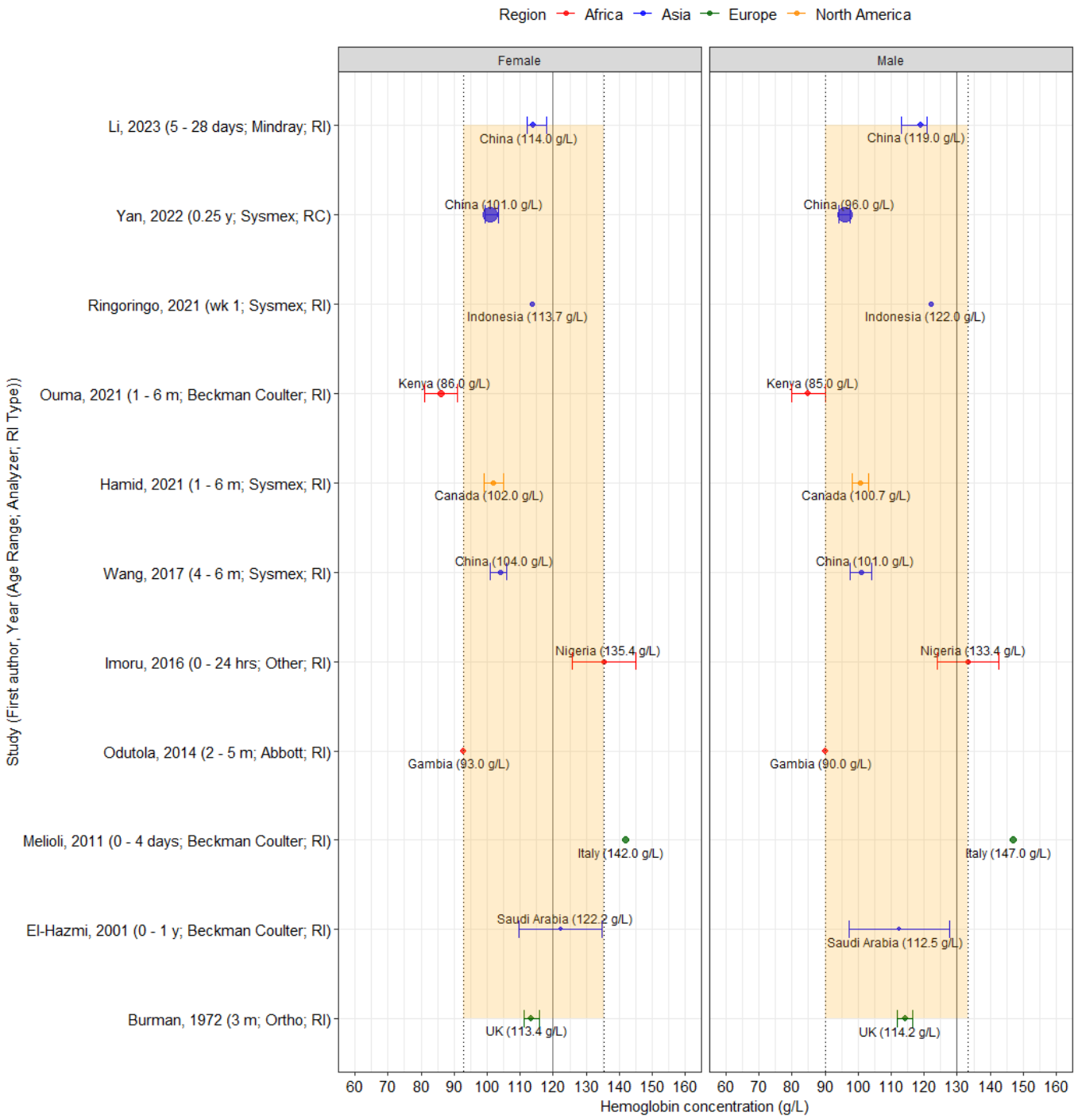


Figure S36. Distribution of lower Hemoglobin RIs limits for 15-18 years old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

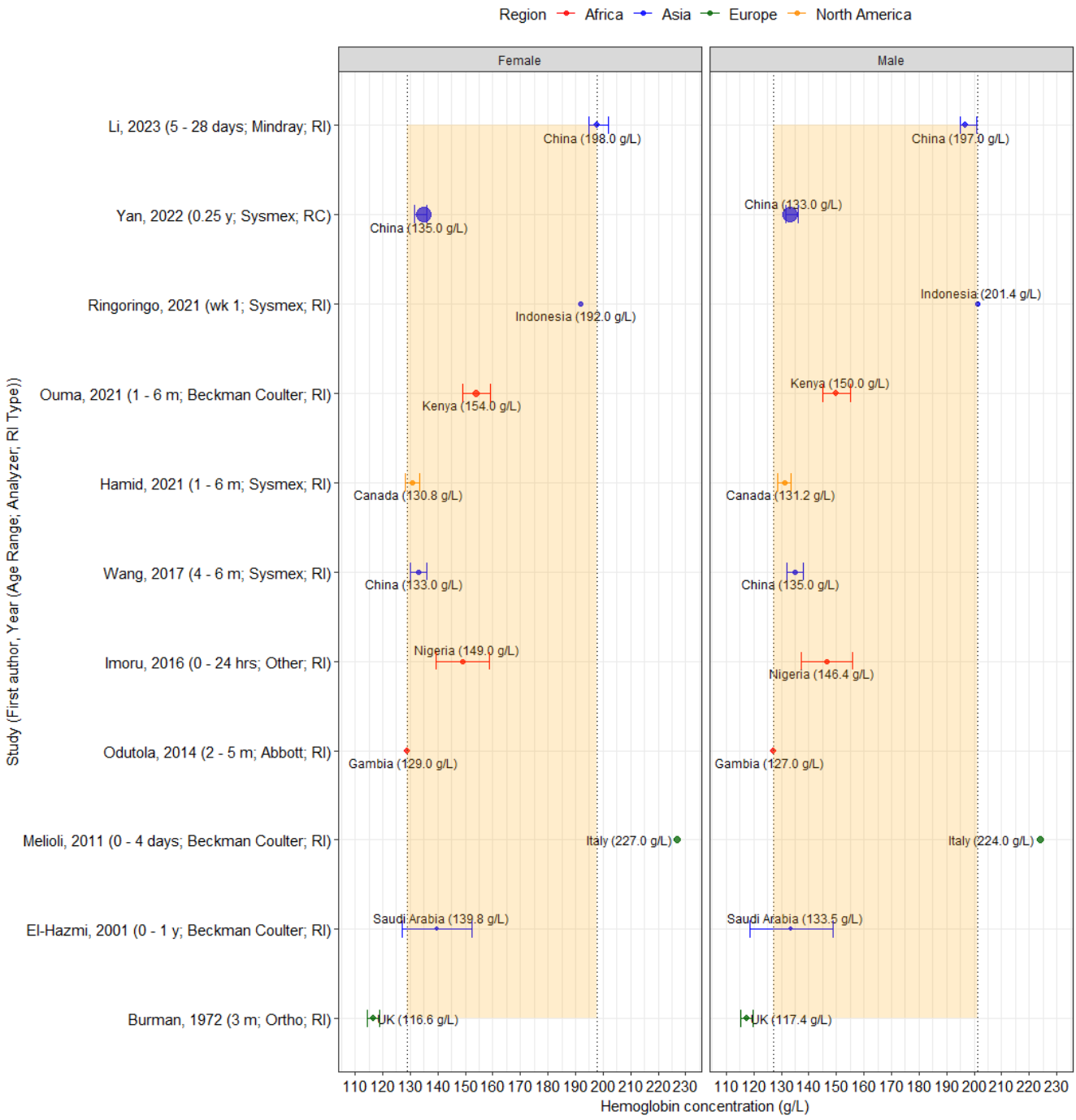


Figure S37. Distribution of upper Hemoglobin RIs limits for 15-18 years old males and females10th and 90th percentile (orange shaded) and WHO anemia threshold for available age ranges (solid vertical line)

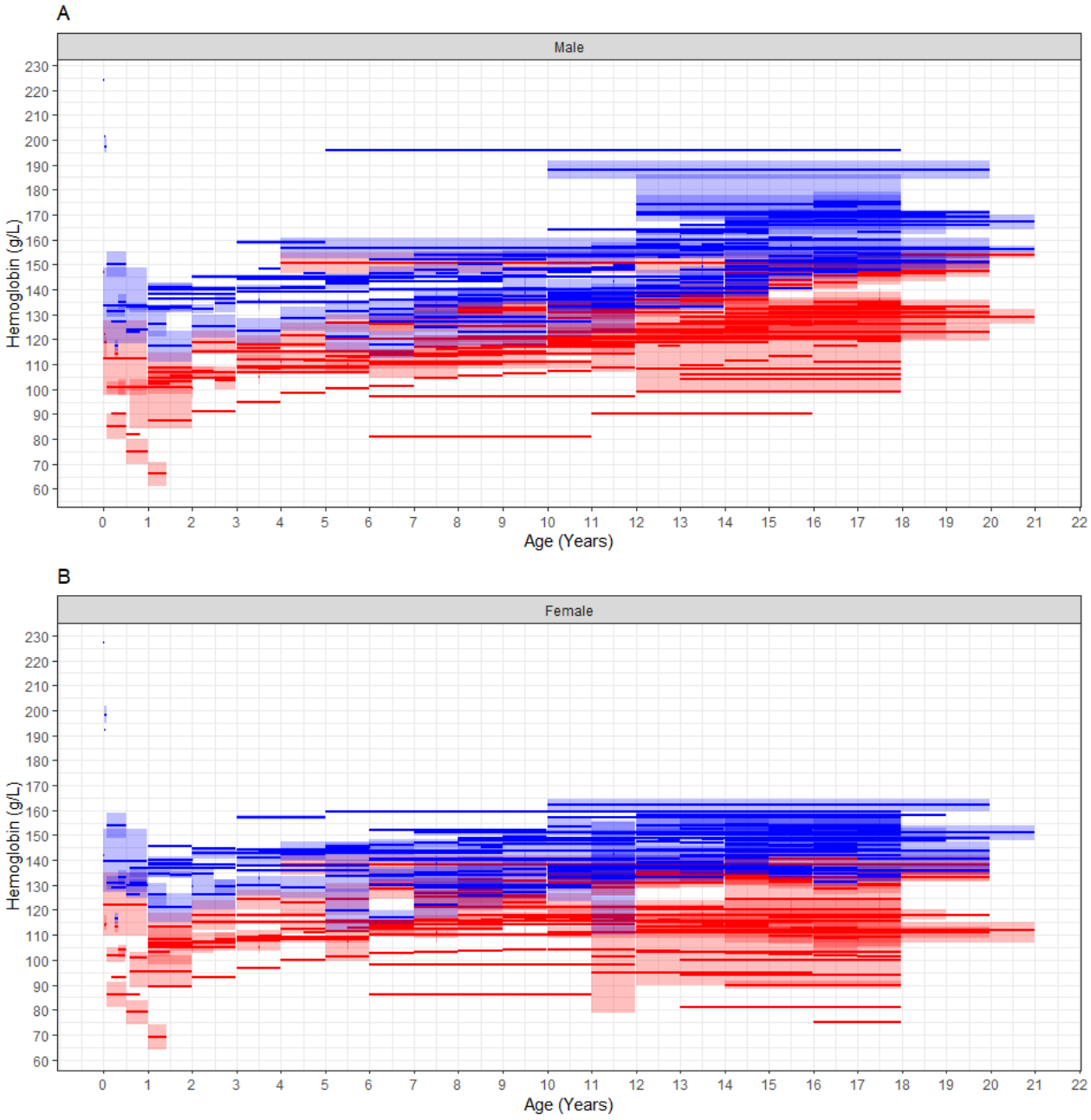


Figure S38. Age specific Hemoglobin reference interval (lower red and upper blue) limits for males (panel A) and females (panel B) with confidence intervals (CIs) from all studies included in this systematic review. Shaded areas represent CIs.

CHAPTER 4: PEDIATRIC REFERENCE INTERVALS AND CURVES FOR FERRITIN ESTIMATED USING DIRECT METHODS: A SYSTEMATIC REVIEW AND META-ANALYSIS

4.1. Preface to Chapter 4

Chapter 4 presents the third manuscript of this thesis, which focuses on evidence synthesis of RIs and RCs for serum and plasma ferritin as biomarkers for iron status in children. Based on methods described in the published protocol in Chapter 2, this manuscript systematically synthesizes the current evidence on pediatric ferritin reference intervals and reference curves, summarizing both lower and upper limits for males and females under 18 years of age. Ferritin has been studied less extensively than hemoglobin. To our knowledge, this is the first systematic review with meta-analysis specifically addressing the pediatric ferritin reference intervals and curves. This is also the first review to separately synthesize lower and upper limits, incorporate both reference intervals and curves, and present the findings with this level of methodological and graphical details. Furthermore, we extend to and tailor the previously developed PRINCES Shiny application to include ferritin specific data and visualization features for interactive exploration of results (Pediatric Reference Intervals and Curves Evidence Synthesis - Ferritin [PRINCES-F]: https://monderic.shinyapps.io/FERRITINprinces/). The manuscript submitted for review to *BMC Pediatrics*.

VB led the study design, screening of references, full text review and selection, data extraction of selected studies, quantitative and narrative data analysis, development of Shiny app, presentation and interpretation of results and has drafted the initial manuscript. MB was involved in screening of references, full text review and selection, data extraction of selected studies, and the review of the manuscript. JL and MP were involved in data extraction from the included studies and revision of the manuscript. BKP, PCP, and JSH were involved in the design, data

interpretation, writing and revision of the manuscript. FM and ML were involved in the design, data interpretation and revision of the manuscript.

4.2 Manuscript status: under review for publication to *BMC Pediatrics*

Pediatric Reference Intervals and Curves for ferritin estimated using direct methods: A Systematic Review and Meta-analysis

Vid Bijelić[1,2], Marijana Bijelić[1], Josh Larock[3], Michael Pham[2,3], Franco Momoli[1], Mira Liebman[2], Beth K. Potter[1], Patricia C. Parkin[4*], Jemila S. Hamid[2,3*]

[1] School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
[2] Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada
[3] Department of Mathematics and Statistics, University of Ottawa, Ottawa, Ontario, Canada
[4] The Hospital for Sick Children, Toronto, Ontario, Canada

[*]Corresponding Authors:
Jemila S. Hamid, MSc, PhD (Statistical Correspondence)
University of Ottawa
STEM Complex, Room# 657
150 Louis-Pasteur Pvt, Ottawa, ON, K1N 6N5
Email: jhamid@uottawa.ca

Patricia C. Parkin, MD, FRCPC (Clinical Correspondence)
The Hospital for Sick Children Research Institute
686 Bay St, Toronto, ON, M5G 0A4, Canada
Email: patricia.parkin@sickkids.ca

Abstract

Introduction: Ferritin plays crucial role in assessing iron status in children. Clinicians rely on reference intervals (RIs) and curves (RCs) that provide lower and upper limits of ferritin in reference population. In contrast, WHO provides thresholds to define iron deficiency or overload based on global expert consensus. Knowledge synthesis provides high-quality information which is often used to inform guidelines and promote evidence-based practice.

Methods: Protocol was published in PROSPERO(CRD42023399802). EMBASE, MEDLINE, SCOPUS, and Cochrane electronic libraries were searched from inception to August 31, 2023. Studies including unhealthy children, missing sex-specific RIs/RCs, or limited to cord blood were excluded. CLSI-adherent studies were meta-analyzed, while remaining eligible studies were narratively synthesized. We pooled sex-specific RI limits with $I^2 < 75\%$. We used forest plots and an R Shiny tool to analyze heterogeneity.

Results: Twenty of 4508 studies were included in systematic review. Sixteen studies involved RIs and four RCs. Eight RIs and one RC were meta-analyzed across standardized age partitions. Lowest pooled RIs limits were observed for 24-33 months old children. Pooled upper RI limit did not exceed maximum observed 95 μg/L at any age. High heterogeneity in upper RI limits was observed for younger females. For many studies, lower RI estimated values were substantially below or above WHO threshold used to define iron deficiency.

Conclusion: For accurate diagnosis and appropriate treatment of children with iron deficiency, laboratories must reconcile RIs, WHO thresholds and other clinical decision limits. There is a need for more rigorous development of RIs from a local reference population. Our study revealed the methodological limitations and the knowledge gaps. We also provided strategies to overcome these challenges, including a novel standardized age partitioning to allow quantitative synthesis using meta-analysis.

Introduction

Most intracellular iron in humans is stored in the ferritin protein.[1] A low concentration of ferritin in serum and plasma has been commonly used over several decades for the diagnosis of iron deficiency.[2] This is preferred to invasive bone marrow aspirate examination.[3] Although less common, a high concentration of ferritin may indicate iron overload, prompting more advanced investigations such as liver magnetic resonance imaging (MRI) or biopsy.[1–3] Detection of low or high ferritin levels may lead to further testing and treatment including iron therapy for iron deficiency or genetic testing for hereditary hemochromatosis for iron overload.[3,4]

Ferritin reference intervals (RIs), representing the lower and upper limits (2.5th percentile and 97.5th percentile) of ferritin in a healthy reference population, are used by clinicians to determine patients' iron status.[2,5] Pediatric RIs need to consider physiologic changes related to healthy child growth and development and thus frequently require extensive age partitioning.[5–7] Reference curves (RCs) that treat age as a continuous variable offer an alternative to such partitioning for Pediatric populations.[8]

The Clinical Laboratory Standards Institute (CLSI) guidelines, an internationally developed set of recommendations for RIs, recommend that clinical laboratories develop their own regional RIs or at the minimum determine if non-regional RIs developed elsewhere are appropriate for their patient population.[5] Previous studies have shown large inconsistencies in the use of published Pediatric RIs across laboratories.[9] This may impact follow-up testing and care: one study found that more than 50% of samples measured at local hospitals would be classified differently (within or outside of the reference interval) depending on which RIs were used,[7,10]

In addition to RIs, clinical decision limits (CDLs) based on clinical outcomes and expert consensus are used for the identification of high-risk conditions using a single threshold.[11] The World Health Organization (WHO) updated their published CDL thresholds to define iron

deficiency in 2020 based on global expert consensus.[3] This creates a challenge for clinicians when reviewing laboratory results and diagnosing patients with iron deficiency: Should they use RIs based on a local reference population or WHO CDLs based on global expert consensus? This is an important health issue, as the WHO reports that iron deficiency affects 43% of children worldwide.[3]

With the objective of filling the knowledge gap related to Pediatric ferritin RIs and providing an improved evidence pool for future recommendations, we conducted a systematic review and synthesized all available evidence on ferritin RIs and RCs for the Pediatric population. We also used narrative synthesis to systematically synthesize evidence from studies all included in the final systematic review. When appropriate, we used meta-analysis to quantitatively synthesize evidence and provided pooled estimates for the lower and upper RI limits. We studied heterogeneity across studies graphically and using formal statistical methods and identified potential sources of heterogeneity in Pediatric ferritin RIs and RCs.

Methods

This systematic review was registered with PROSPERO (CRD42023399802) and the protocol was published.[12] We report the results according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[13]

Search, screening, and data extraction

The search strategy was developed and implemented in consultation with a research librarian, which involved a combination of words and MeSH terms specific to MEDLINE, EMBASE, SCOPUS, and the Cochrane Library published from inception to the search date (August 31, 2023). VB and MB independently screened citations retrieved by the search against eligibility criteria in each of two stages: titles and abstracts were screened in stage one and full-text articles for citations passing the first stage were screened in stage two. Data was extracted by JL and MP

and verified by VB. Studies that included a healthy Pediatric population, younger than <18 years of age, provided RIs and/or RCs for ferritin, and published in English were included. We extracted information on lower and upper RI limits for each age and sex partition (as provided in the studies) with corresponding confidence intervals (CIs). For studies with missing CIs, the Solberg formula[14] was used to calculate CIs, provided that the studies used the parametric RI method and provided standard deviation and sample size. For RCs, if a regression equation was provided, the mid-point of a pre-specified age partition was used to calculate RI limits. Studies that did not report male or female RIs and RCs separately, and studies that reported RIs and RCs obtained only from cord-blood were excluded. In situations where multiple studies used the same database or cohort, only the study with the largest sample size, or the most recent study if sample sizes were identical, was included.

Risk of Bias Assessment and confidence in cumulative evidence

There is no risk of bias (RoB) tool that assesses the quality of studies involving RIs and RCs; thus, formal RoB assessment was not possible. To overcome this challenge, we examined whether or not the studies adhered to the CLSI guidelines[5], as this guideline is a gold standard for establishing RIs. The CLSI guidelines reflect long-established best practices in four key areas: outlier detection, partitioning (finding homogeneous age subgroups), RI estimation methods, and estimation of the corresponding confidence intervals for the upper and lower RI limits.[5]

This systematic review synthesized studies that estimate RIs and RCs, and not the studies that test hypothesized associations between variables. Thus, publication bias is not relevant to this systematic review.[15,16] To our knowledge, The Grading of Recommendations Assessment, Development and Evaluation (GRADE)[17,18] adaptation nor any other tool exist for assessing confidence in cumulative evidence for studies estimating RIs and/or RCs, hence we described

quality and strength of evidence narratively using the CLSI guidelines and other available best practices in terms of RI and RC estimation.

Evidence Synthesis

We used quantitative synthesis using meta-analysis as well as narrative synthesis. Only RIs and RCs from studies that adhered to the CLSI guidelines were included in the meta-analysis. Different age partitions were used across studies; therefore, we created standardized age partitions, informed by literature and clinical practice. For children below 3 years, we used 3-month age ranges; for children above 3 years, we used 1-year age ranges. To study age partitioning as a source of heterogeneity in the evidence synthesis, the age partitions used in each study were categorized as "within range" (study age partition corresponded to or was fully embedded within our standardized age partitions) or "out of range" (study age partition was wider than standardized age partitions). Meta-analyses within each age partition were performed separately for males and females, and for lower and upper RI limits. Heterogeneity was assessed graphically using forest plots and quantitatively using the I^2 statistic. We also presented the I^2 statistic graphically using heatmaps. The pooled estimates for ferritin RIs were calculated only when $I^2 \leq 75\%$, as recommended by the Cochrane Handbook for Systematic Reviews.[19] To account for heterogeneity, random effects meta-analysis was used.

We were unable to include many studies in the meta-analysis due to heterogeneity and lack of adherence to the CLSI guidelines. Therefore, we also conducted a narrative synthesis of all included studies, using both qualitative and quantitative methods. In our narrative synthesis, we examined the distribution of the lower and upper RI limits across all included studies in relation to the recently published age-specific WHO thresholds for iron deficiency. We also developed a novel web-based graphical and computational tool (entitled: Pediatric Reference Intervals and Curves: Evidence Synthesis for Ferritin [PRINCES-FERRITIN

https://monderic.shinyapps.io/FERRITINprinces/). The tool enables users to specify combinations of factors that might be considered as potential sources of heterogeneity, including age category, region, and methodological quality. It allows for the exploration of the impact of these factors, either individually or in combination, on the variability of the lower and upper RI limits for both males and females. All analyses were performed using the R statistical software version 4.3.1, the *metagen* function from *meta* package version 7.0-0 and *metafor* package version 4.6-0 were used for meta-analysis, and the web-based tool was developed using R Shiny version 1.9.1.[20,21]

Patient and public involvement

No patients or members of the public were involved in this systematic review.

Deviations from the published protocol

In the development of our protocol, we planned to use the Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2)[22] to assess RoB. However, after further evaluation QUADAS-2 was deemed not suitable for assessing RoB of RIs and RCs. As such, we used the CLSI guidelines to examine RoB and assess evidence quality and strength.

Results

The search yielded 4508 citations of which 3040 were duplicates. A total of 1468 titles and abstracts were screened at the first stage, leading to 46 studies to be retained for full text review. After the full text review, 20 studies (involving 26,006 children)[23–42] met all eligibility criteria for inclusion in the review (Figure 1, Table 1). Summary statistics are provided in Table 1 and a more detailed description of the included studies is provided in the supplementary file (Table S1). Of the 26, 006 children from which evidence was synthesized, 12,863 (49.5%) were girls and 13,143 (50.5%) were boys. Eight studies (involving 7,604 children) reporting RIs that adhered to CLSI guidelines were included in the meta-analysis, where pooling was done where

appropriate. One RC study (involving 890 children) provided 90% CIs, hence was also included in the quantitative synthesis. The narrative synthesis and the graphical examination included all 20 studies.

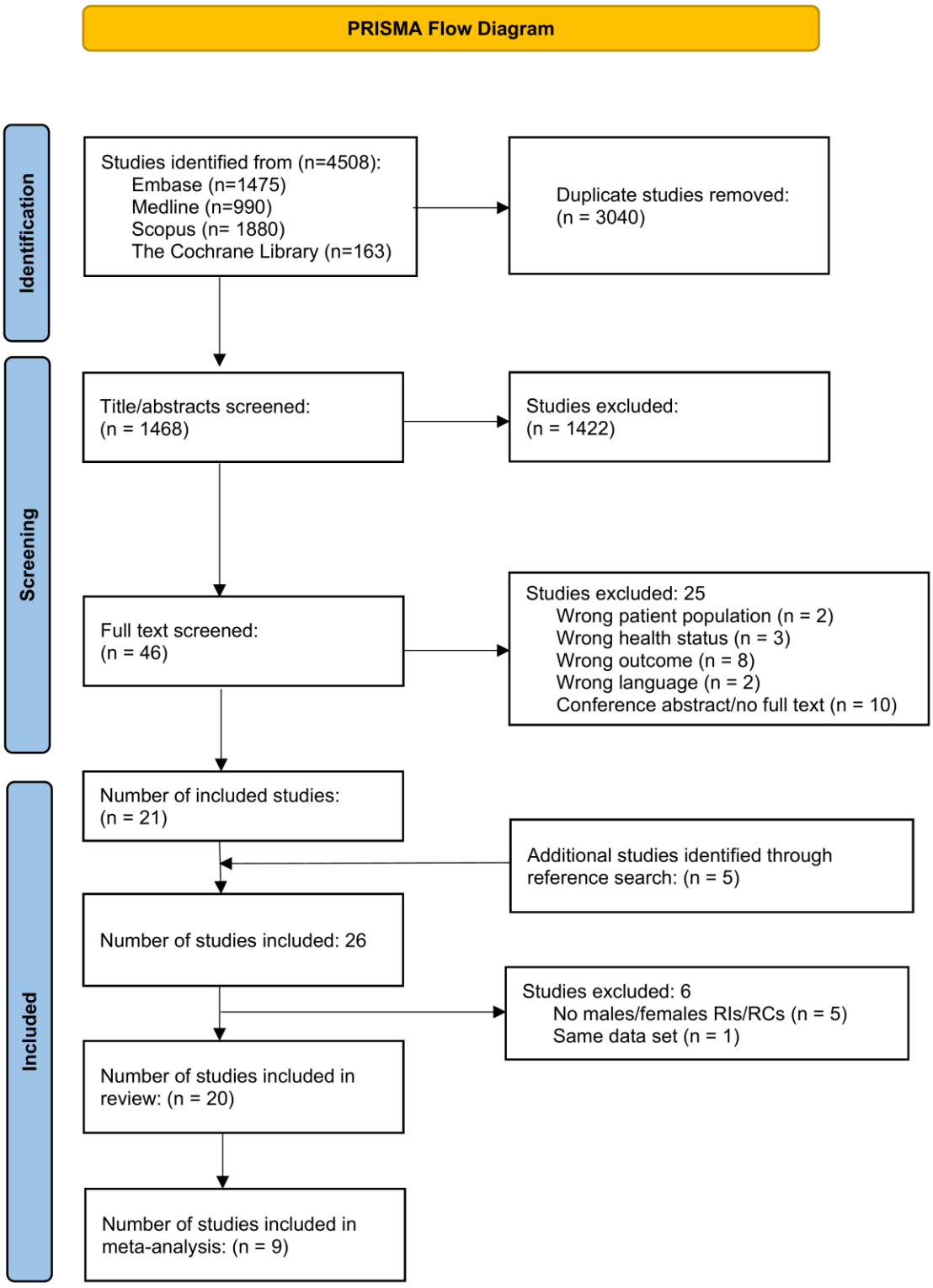


Figure 1. The PRISMA Diagram outlining the search and screening process of the systematic review (Adapted from: Page et al. 2020)

The publication year for included studies ranged from 1976 to 2023 and their geographical distribution covered North America (n=12, 60%), Europe (n=6, 30%), and Asia (n=2, 10%). Sixteen (80%) studies provided RIs and four (20%) provided RCs (Table 1, Table S1). The most commonly used analyzers were Abbott (30%), Roche (30%) and Siemens (20%) models.

Table 1. Characteristics of studies included in the systematic review, including analytical and methodological considerations for establishing RIs and RCs, and study adherence to the CLSI guidelines

Characteristics	**n (%)**	**Number of Children**
Total	20	26006
Sex		
Males	20 (100.0)	13143 (50.5)
Females	19 (95.0)	12863 (49.5)
Regions		
Australia & Papua New Guinea	2 (10.0)	1366 (5.3)
Europe	6 (30.0)	4163 (16.0)
North America	12 (60.0)	20477 (78.7)
Publication Years (min - max)	1976 - 2023	-
Interval Type		
Reference Intervals	16 (80.0)	21457 (82.5)
Reference Curves	4 (20.0)	4549 (17.5)
Estimation method		
Reference Intervals		
Non-parametric, n (%)	9 (56.2)	18323 (85.4)
Parametric, n (%)	4 (25.0)	2228 (10.4)
Robust, n (%)	6 (37.5)	804 (3.7)
Other (Method), n (%)	1 (6.2)	102 (0.5)
Reference Curves		
LMS, n (%)	2 (50.0)	3279 (72.1)
Quantile regression, n (%)	2 (50.0)	1270 (27.9)
Age (min - max)	0 - 24.99	-
Analyzer models		
Abbott	6 (30.0)	3534 (12.8)
Beckman Coulter	3 (15.0)	10743 (39.0)
Ortho	2 (10.0)	577 (2.1)
Roche	6 (30.0)	8242 (29.9)
Siemens	4 (20.0)	2583 (9.4)
Other	1 (5.0)	753 (2.7)
Not provided	2 (10.0)	1094 (4.0)
Adhered to CLSI Guidelines**	8 (50.0)	7604 (35.4)
Outlier detection, n (%)	10 (62.5)	8692 (40.5)
Partitioning, n (%)	15 (93.8)	21253 (99)
Method for RIs estimation, n (%)	14 (87.5)	20787 (96.9)
Estimation of CIs for RIs	12 (75.0)	19015 (88.6)

** Does not always sum to 100% due to studies in some cases involving more than one option*

*** Only applicable to studies reporting RIs (n=16 studies, n=21457 children)*

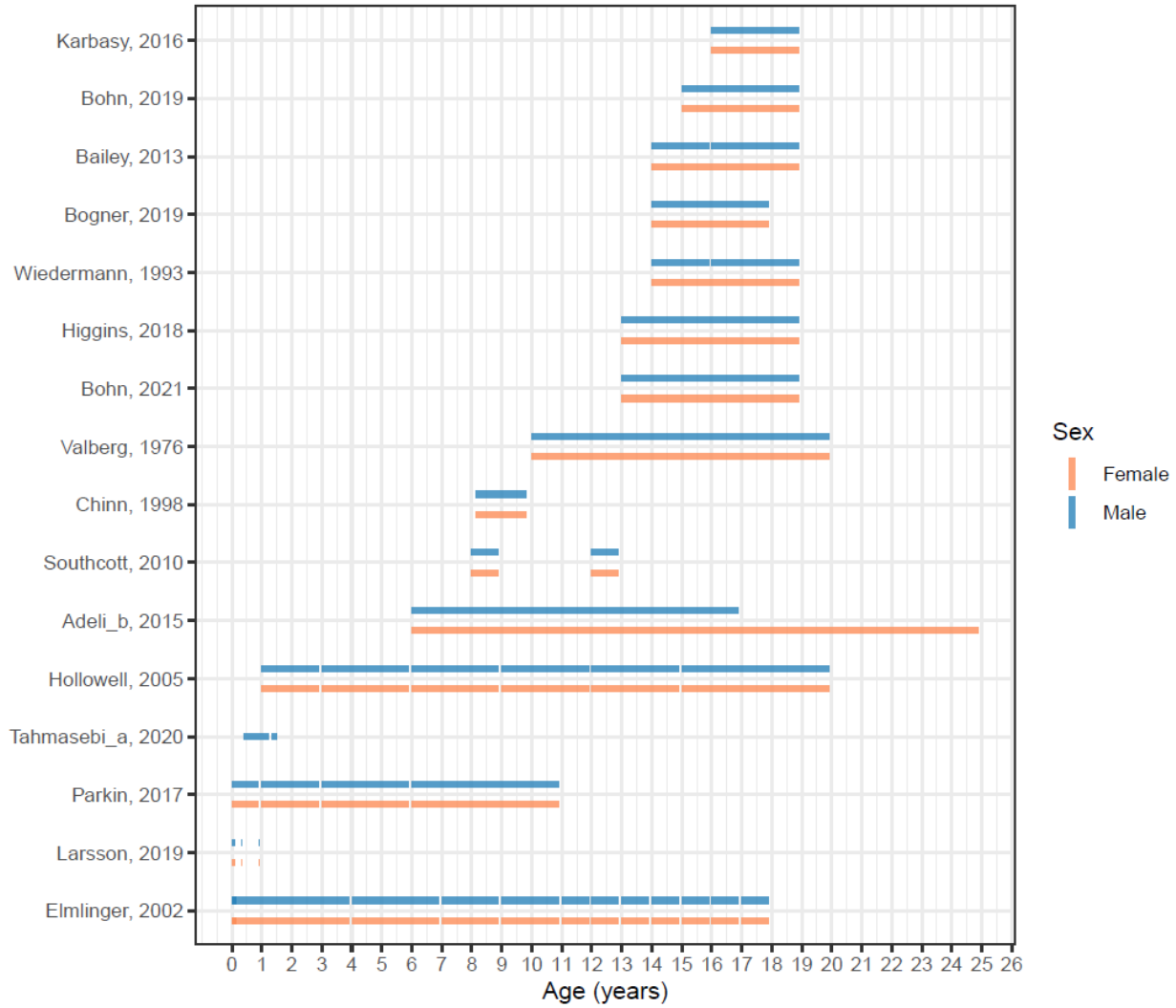


Figure 2. Interval width of the age partitioning used when establishing ferritin RIs presented by sex. Studies were ordered (bottom to top) according to the minimum age for the study (reference) population

Figure 2 shows inconsistencies in age partitioning across twenty studies, for both sexes. The length of age intervals varied from as narrow as less than one week (e.g., 48 to 72 hours) to as wide as 19 years (e.g., 6 years to 25 years).

Evidence Synthesis

Meta-Analysis

Eight RI studies[23,24,26,27,31,34,36,39] followed all four key recommendations from the CLSI guidelines[5] and one RC study[30] provided CIs for lower and upper limits. Hence, nine (45%) of twenty studies consisting of 8,494 children were included in the meta-analysis. The most common reasons for non-adherence to the CLSI were lack of outlier detection (n=6, 37.5%) and calculations/reporting of CIs for RIs (n=7, 35.0%), followed by the choice of statistical methods for estimation of RIs (n=2, 12.5%), and related to data partitioning (n=1, 6.3%) (Table 1, Table S2). The pooled RI limits are provided in Table 2 and forest plots corresponding to the meta-analysis (leading to these pooled estimates) are presented in Figure 4 and supplemental material section (Figures S1-S16).

Table 2. The pooled estimates of the lower and upper RI limits for ferritin (ug/L) obtained from random effects meta-analysis. The 16 studies adhering to the CLSI guidelines were included in the analysis and pooling was only done when $I^2 \leq 75\%$. Standardized age partitions were used, and analysis was performed for males and females separately.

	Pooled Lower RI limit RIs (90% CI)		**Pooled Upper RI limit** RIs (90% CI)	
Age partitions*	**Males**	**Females**	**Males**	**Females**
0-3 months	**	**	**	**
3-6 months	9.1 (8.5, 9.7)	**	**	**
6-9 months	10.4 (8.0, 12.9)	**	95.7 (71.8, 119.6)	**
9-12 months	9.7 (6.8, 12.5)	**	**	**
12-15 months	**	**	79.6 (56.7, 102.6)	82.4 (76.5, 88.3)
15-18 months	**	**	**	**
18-21 months	7.5 (4.8, 10.2)	7.9 (4.5, 11.3)	**	78.4 (72.8, 83.9)
21-24 months	**	**	**	**
24-27 months	6.0 (5.5, 6.5)	7.3 (4.7, 10.0)	71.4 (65.5, 77.3)	**
27-30 months	**	**	**	**
30-33 months	6.0 (5.5, 6.5)	7.0 (4.3, 9.7)	70.4 (64.6, 76.2)	**
33-36 months	**	**	**	**
3-4 years	12.0 (11.5, 12.5)	**	70.6 (67.7, 73.5)	**
4-5 years	12.0 (11.5, 12.5)	**	70.3 (67.4, 73.1)	**
5-6 years	12.0 (11.5, 12.5)	11.0 (8.0, 14.0)	66.7 (46.9, 86.5)	**
6-7 years	15.0 (14.1, 16.0)	11.2 (8.5, 13.9)	**	**
7-8 years	15.1 (14.1, 16.0)	11.6 (8.9, 14.4)	**	**
8-9 years	15.1 (14.2, 16.0)	12.0 (9.3, 14.8)	**	81.0 (70.9, 91.1)
9-10 years	15.2 (14.3, 16.1)	12.5 (9.8, 15.2)	**	82.1 (71.8, 92.5)
10-11 years	15.4 (14.4, 16.1)	**	**	83.0 (72.2, 93.8)
11-12 years	15.6 (14.0, 17.1)	**	**	79.7 (66.2, 93.1)
12-13 years	15.2 (13.3, 17.1)	10.7 (8.2, 13.1)	**	80.8 (70.6, 89.4)
13-14 years	12.9 (11.6, 14.3)	8.2 (6.9, 9.6)	**	76.4 (69.4, 83.4)
14-15 years	12.6 (11.3, 13.9)	**	**	70.3 (67.1, 79.9)
15-16 years	13.9 (11.2, 16.6)	**	**	**
16-17 years	14.5 (11.8, 17.1)	**	**	**
17-18 years	**	**	**	**

**Represents the standardized age-partitions we used in our systematic review*
*** Pooling was not possible because heterogeneity I^2 was > 75%*

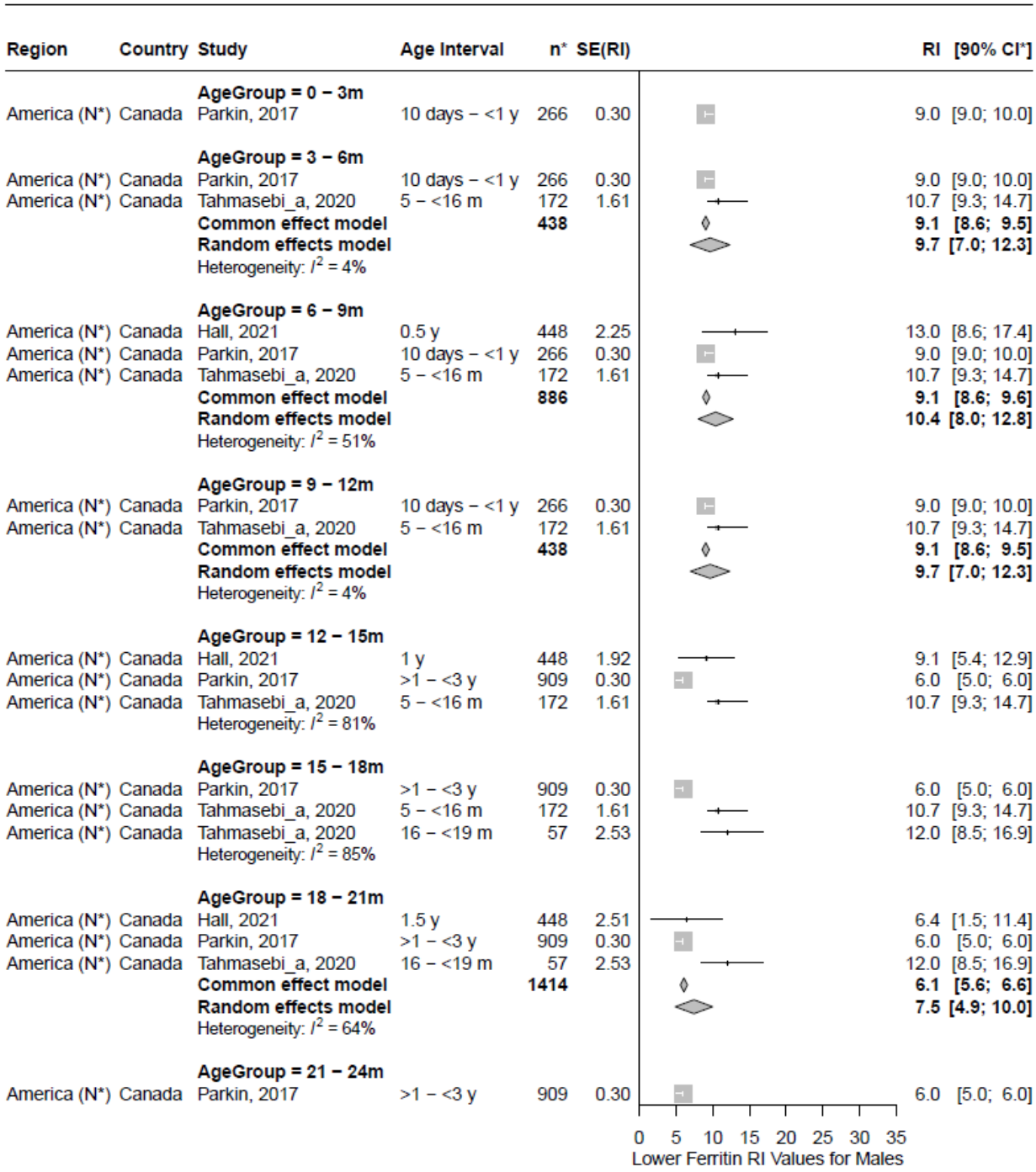


Figure 3. Example forest plot of ferritin RIs estimates for males by standardized age partitions (0–24 months). Studies are ordered by region, country, and first author's last name. Additional forest plots are available in the supplemental material section. N* means North

Pooled estimates were provided for a limited number of age partitions due to high heterogeneity (Table 2, Figures S1-S16). The data for 0-3 months age group was exceedingly heterogeneous to generate a pooled estimate for this age group (Figure 3; Figures S1-S2). The pooled lower RI limit showed a decrease from 3 months to 12 years of age, reaching the lowest value of 6.0 μg/L between the age of 24 and 33 months (Table 2). From 3 years to pre-adolescence, the lower RI limit increased to approximately 12 μg/L (females) and 15 μg/L (males), followed by a low value of 8 μg/L for adolescent females. The highest pooled upper RI limit was 95.7 μg/L, recorded for age 6-9 months (Table 2; Figures S1-S16).

Narrative Synthesis

Detailed descriptions of the 20 studies are presented in the supplementary file (Table S1). Variation of RIs across different age categories, as reported by the studies, is provided in Figure 4, where lower and upper reference interval limits are plotted against age. The results show that the RIs indeed vary across age categories, especially for younger children.

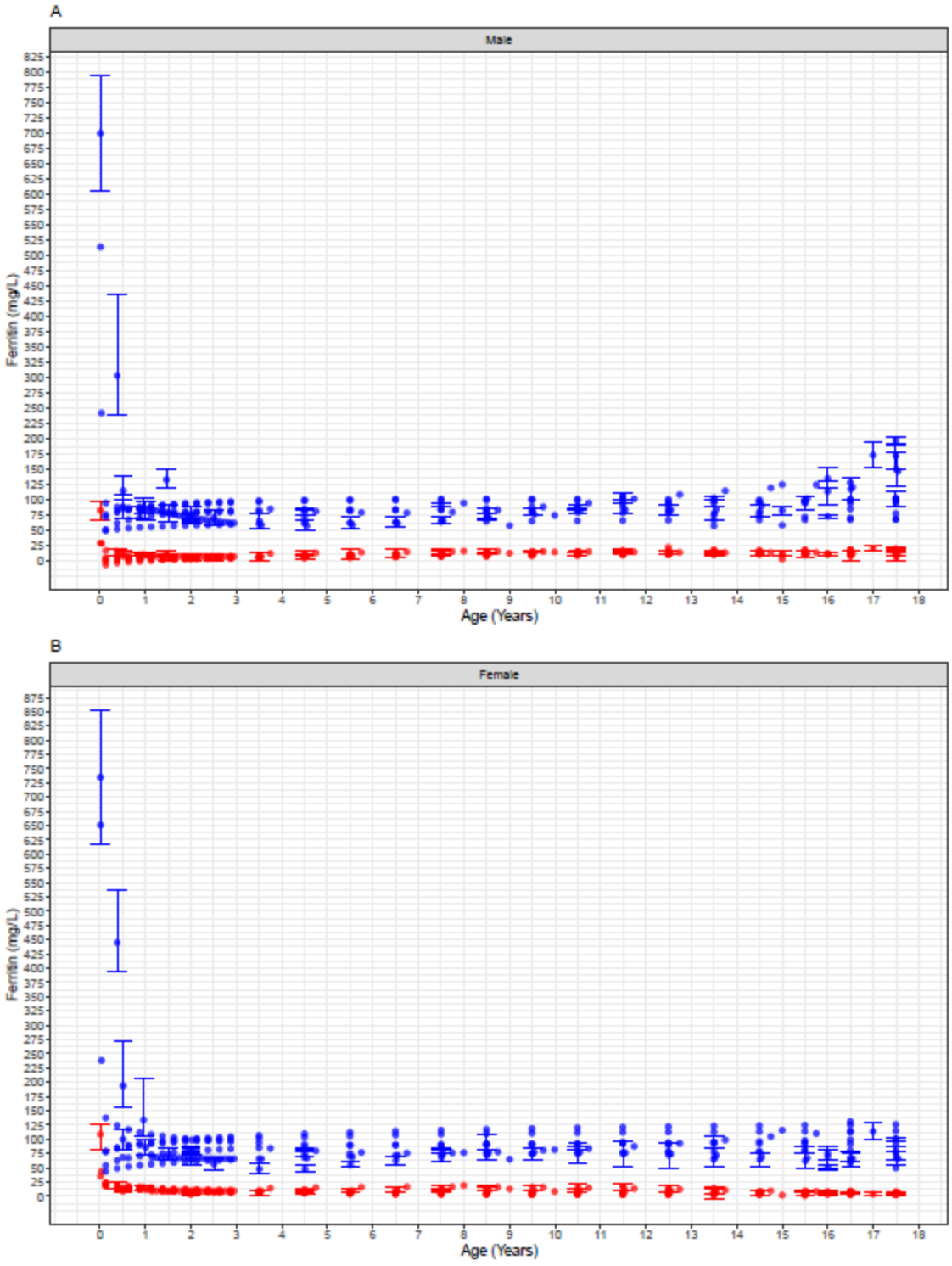


Figure 4. Age-specific ferritin RIs (lower limits in red and upper limits blue) for males (panel A) and females (panel B) with corresponding confidence intervals (CIs), where available, from all studies included in this systematic review

A more comprehensive and in-depth assessment of both the lower and upper RI limits across all age partitions and separately for males and females are depicted in narrative synthesis figures (Figures S18-S25). We also explored regional differences in the narrative synthesis. Our novel web-based graphical and computational tool (PRINCES-FERRITIN) is available online (https://monderic.shinyapps.io/FERRITINprinces/) and gives a more comprehensive look at the distribution of Pediatric RIs and RCs for ferritin and explores differences in respect to additional factors and facilitates further interpretation of the of the narrative synthesis results provided in this paper. Supplementary material section (Figure S17) offers a snapshot of one scenario captured from our web-based tool.

Children 0 – 23 months of age

For children 0 – 23 months (Table 3; Figures S18-S19), the lower RI limits for males ranged from 0.71 ug/L[33] to 83.00 ug/L[35] and for females from 3.00 ug/L[32] to 109.00 ug/L[35], reported in the first 48-72 hours of life[35]. The upper RI limits ranged from 49.82 ug/L (males) and 43.92 ug/L (females)[33] to 699.00 ug/L (males) and 735.00 ug/L (females).[35]

Table 3. Descriptive statistics of lower and upper ferritin limits by WHO age partitions

Age Partition	Lower Reference Limits			Upper Reference Limits		
	Median (IQR)	Range (Min, Max)	Mean (SD)	Median (IQR)	Range (Min, Max)	Mean (SD)
Females						
0-23 months	11.9 (9.8, 16.0)	(3.0, 109.0)	15.2 (13.6)	82.7 (67.7, 99.6)	(43.9, 735.0)	112.1 (120.6)
24-59 months	8.7 (7.4, 9.5)	(3.0, 15.4)	8.7 (2.1)	81.0 (66.2, 97.9)	(48.0, 109.8)	80.7 (16.7)
5-10 years	8.4 (4.6, 12.0)	(3.0, 19.1)	8.8 (4.5)	81.9 (74.6, 91.5)	(55.2, 119.0)	85.5 (15.7)
10-18 years	6.0 (4.0, 8.5)	(1.6, 17.5)	6.8 (3.7)	92.0 (75.7, 104.9)	(49.9, 131.2)	90.4 (19.7)
Males						
0-23 months	6.0 (4.8, 8.5)	(0.7, 83.0)	8.7 (11.2)	81.5 (72.7, 88.6)	(49.8, 699.0)	101.4 (100.0)
24-59 months	6.8 (5.2, 7.7)	(3.0, 13.4)	6.9 (2.1)	80.0 (65.5, 84.2)	(57.5, 99.9)	77.1 (13.2)
5-10 years	11.5 (8.7, 14.2)	(5.7, 16.40	11.4 (3.0)	82.5 (70.4, 89.9)	(58.0, 106.7)	81.5 (13.6)
10-18 years	13.2 (11.5, 15.3)	(3.0, 22.7)	13.5 (3.2)	94.9 (79.4, 101.6)	(57.5, 197.2)	95.4 (24.6)

Children 24 – 59 months of age

For children 24 - 59 months (Table 3; Figures S20-S21), lower RI limits for males ranged from 3.00 ug/L[32] to 13.40 ug/L[37] and for females from 3.00 ug/L[32] to 15.40 ug/L[37]. The upper RI

limits for males ranged from 57.50 ug/L[42] to 99.92 ug/L[33] and for females from 48.00 ug/L[30] and 109.84 ug/L[33].

Children 5 – 10 years of age

For children 5 – 10 years (Table 3; Figures S22-S23), the lower RI limits for males ranged from 5.73 ug/L[33] to 16.40 ug/L[29] and for females from 3.01 ug/L to 19.14 ug/L[29]. The upper RI limits for males ranged from 58.00 ug/L[28] to 106.70 ug/L[23] and for females ranged from 55.20 ug/L[30] to 119.01 ug/L[33].

Children 10 –18 years of age

For children 10 – 18 years (Table 3; Figures S24-S25), the lower RI limits for males ranged from 3.00 ug/L[40] to 22.74 ug/L[29] and for females from 1.60 ug/L[29] to 17.46 ug/L[30]. The upper RI limits for males ranged form 57.45 ug/L[29] to 197.20 ug/L[32] and for females from 49.89 ug/L to 131.20 ug/L[29].

Substantial variability related to age, sex, and country was observed (Figures S18-S26; PRINCES-FERRITIN-web-based tool). The lower RI was substantially different (below or above) than the WHO threshold (12 µg/L for children under 5 years old and 15 µg/L for older children) to define iron deficiency.

There also appears to be regional variation across the different continents. For many studies, the lower RI was substantially different (lower or higher) than the WHO threshold to define iron deficiency. However, this is to be expected considering that WHO thresholds are calculated using 5th percentiles of the population, whereas lower limits of RIs correspond to the 2.5th percentile of the population.

Discussion

This systematic review of Pediatric RIs and RCs for ferritin includes 20 studies published in English from 1976 to 2023 across Australia, Europe, and North America. Only eight Pediatric

studies involving ferritin RIs followed the CLSI guidelines and one study reporting RCs was included in the meta-analysis. All 20 studies were included in the narrative synthesis. Our novel web-based PRINCES- FERRITIN visualization tool presents all data discussed in the current work alongside the WHO iron deficiency thresholds. None of the 20 studies excluded ferritin data from participants with laboratory evidence of inflammation or anemia, or clinical risk factors for iron deficiency. Thus, with respect to optimal iron health, the study reference populations may not truly represent 'healthy' populations.

The pooled lower RI limit reached the lowest values in young males (6 μg/L) and females (7 μg/L) children 24-33 months of age and adolescent females (8 μg/L). The pooled upper RI limit did not exceed maximum observed 95 μg/L at any age. Substantially higher heterogeneity was observed for younger compared to older children, females compared to males, and upper compared to lower RI limits. The narrative synthesis showed that for many studies, the lower RI was substantially different (lower or higher) than the WHO threshold to define iron deficiency.

There is a growing body of literature aimed at determining a CDL for iron deficiency based on physiologic or clinical outcomes in Pediatric populations with high prevalence of iron deficiency: young preschool age children and adolescent females.[43–46] Several investigators have examined the relationship between ferritin and hemoglobin to define a threshold for iron deficiency in young children.[43–45] Abdullah et al examined data from children 12-36 months participating in a Canadian research cohort (excluding data from children with high CRP) and identified a ferritin threshold of 18-24 μg/L.[43] Mei et al examined data from children 12-59 months using the U.S. National Health and Nutrition Examination Survey (NHANES) and identified a ferritin threshold of about 20 μg/L.[44] Mukhtarova and colleagues examined data from children 9-13 months undergoing screening in U.S. outpatient family medicine and

Pediatric clinics and identified a ferritin threshold as high as 24-25 μg/L.[45] Parkin et al examined the relationship between ferritin and cognitive function in children 12-40 months and identified a ferritin threshold of 17 μg/L.[46]

For adolescent females, findings from studies on adults is informative. Mei et al examined data from non-pregnant women 15-49 years (excluding data from individuals with high CRP) using the U.S. National Health and Nutrition Examination Survey (NHANES) and identified a ferritin threshold of 25 μg/L.[44] Truong et al conducted a systematic review of ferritin RIs in adults and identified a median lower RI limit of 8 μg/L in females and advocated for the use of CDLs for the diagnosis of iron deficiency.[47]

Clinicians and laboratories are now confronted with widely different thresholds for interpreting ferritin results when considering the diagnosis of iron deficiency. This is especially true for Pediatric populations at highest risk. For young preschool age children, the lower RI limit (according to our systematic review and meta-analysis) may be as low as 6 μg/L; the WHO threshold (derived by consensus) is 12 μg/L; and the CDL (using physiologic or clinical outcomes) may be as high as 25 μg/L. For adolescent females, the lower RI limit (according to our systematic review and meta-analysis) may be as low as 8 μg/L; the WHO threshold (derived by consensus) is 15 μg/L; and the CDL (using physiologic or clinical outcomes) may be as high as 25 μg/L. To address this, in July 2024 the Ontario Association of Medical Laboratories published new guidelines on the interpretation of ferritin for iron deficiency, recommending a CDL of < 30 μg/L in adults and < 20 μg/L in children which have been adopted by community laboratories.[48]

This systematic review represents the first attempt to provide analysis of published ferritin Pediatric lower and upper RIs and RCs, utilizing both meta-analysis and narrative

synthesis. In this study the effort was made to compare WHO cut-off points with published literature and present findings in a novel and user-friendly web-based tool. Nevertheless, this study also has several limitations. Formal risk of bias assessment was not possible because there is no risk of bias tool for the assessment of studies involving RIs and RCs. However, we attempted to identify higher quality studies by examining adherence to the CLSI guidelines, considered internationally as a gold standard for establishing RIs. Also, our review was limited to studies published in English. This may have led to exclusion of important studies published in other languages. High heterogeneity and low number of studies in a number of standardized age partitions precluded us from providing pooled estimates for those partitions.

There are several implications for clinicians, laboratories and future research. Clinicians must recognize that RIs from their local reference population may not represent optimal iron status; for example, using the ferritin lower RI limit may lead to underdiagnosis and undertreatment when considering WHO defined cut-off points to define iron deficiency. Laboratories must reconcile the various thresholds derived using different approaches. For example, for the diagnosis of iron deficiency, laboratories could provide both RIs from a local reference population as well as a single CDL derived from consensus or based on physiological or clinical outcomes. For future research, our evidence synthesis of ferritin RIs/RCs derived from local reference populations revealed substantial heterogeneity and varying age partitions. These findings suggest the need for more rigorous development of RIs ensuring clinically meaningful age and sex partitions that follow developmental variations, and enrollment of healthy participants with optimal iron status, where the health status of participants is investigated with respect to known risk factors. Further, development of RCs may eliminate the need for extensive age partitioning and may be accompanied by percentile curves and charts similar to WHO

growth charts. Curve estimations are valid both for estimating RIs as well as thresholds and clinical decision limits, where continuous curves representing different percentiles of the population can be provided as a continuous function of age.

Acknowledgement

The authors acknowledge the support from the University of Ottawa in the library search.

Funding

This research is supported by JSH's discovery grant from the Natural Sciences and Engineering Research Council of Canada, Government of Canada (Grant number: RGPIN-2024-05385). The funder had no role in the design and development of this protocol. Authors have no conflicts of interest to declare.

References

1 Knovich MA, Storey JA, Coffman LG, Torti SV, Torti FM. Ferritin for the clinician. *Blood Rev* 2009; **23**: 95–104.

2 Baker RD, Greer FR, The Committee on Nutrition. Diagnosis and Prevention of Iron Deficiency and Iron-Deficiency Anemia in Infants and Young Children (0–3 Years of Age). *Pediatrics* 2010; **126**: 1040–50.

3 World Health Organization. WHO guideline on use of ferritin concentrations to assess iron status in individuals and populations. Geneva: World Health Organization, 2020 https://iris.who.int/handle/10665/331505 (accessed Aug 5, 2024).

4 Speckert M, Ramic L, Mitsakakis N, Bijelić V, Liebman M, Leung E. Severe iron deficiency anemia in the paediatric emergency department: A retrospective study. *Paediatr Child Health* 2023; **28**: 30–6.

5 Clinical and Laboratory Standards Institute (CLSI), editor. Defining, Establishing and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition. Wayne, Pa: Clinical and Laboratory Standards Institute, 2008.

6 Daly CH, Higgins V, Adeli K, Grey VL, Hamid JS. Reference interval estimation: Methodological comparison using extensive simulations and empirical data. *Clin Biochem* 2017; **50**: 1145–58.

7 Lyle AN, Pokuah F, Dietzen DJ, *et al.* Current State of Pediatric Reference Intervals and the Importance of Correctly Describing the Biochemistry of Child Development: A Review. *JAMA Pediatr* 2022; **176**: 699.

8 Higgins V, Adeli K. Advances in pediatric reference intervals: from discrete to continuous. *J Lab Precis Med* 2018; **3**: 3–3.

9 Adeli K, Higgins V, Seccombe D, *et al.* National Survey of Adult and Pediatric Reference Intervals in Clinical Laboratories across Canada: A Report of the CSCC Working Group on Reference Interval Harmonization. *Clin Biochem* 2017; **50**: 925–35.

10 Alnor AB, Vinholt PJ. Paediatric reference intervals are heterogeneous and differ considerably in the classification of healthy paediatric blood samples. *Eur J Pediatr* 2019; **178**: 963–71.

11 Ozarda Y, Sikaris K, Streichert T, Macri J, on behalf of IFCC Committee on Reference intervals and Decision Limits (C-RIDL). Distinguishing reference intervals and clinical decision limits – A review by the IFCC Committee on Reference Intervals and Decision Limits. *Crit Rev Clin Lab Sci* 2018; **55**: 420–31.

12 Bijelić V, Potter B, Parkin PC, Momoli F, Liebman M, Hamid JS. Paediatric reference intervals and curves for haemoglobin and ferritin: protocol for a systematic review and meta-analysis. *BMJ Open* 2023; **13**: e073783.

13 Page MJ, McKenzie JE, Bossuyt PM, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; **372**: n71.

14 Solberg HE. Approved recommendation (1987) on the theory of reference values. Part 5. Statistical treatment of collected reference values. Determination of reference limits. *Clin Chim Acta* 1987; **170**: S13–32.

15 Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629–34.

16 Lin L, Chu H. Quantifying Publication Bias in Meta-Analysis. *Biometrics* 2018; **74**: 785–94.

17 Iorio A, Spencer FA, Falavigna M, *et al.* Use of GRADE for assessment of evidence about prognosis: rating confidence in estimates of event rates in broad categories of patients. *BMJ* 2015; **350**: h870–h870.

18 Holger Schünemann, Jan Brożek, Gordon Guyatt, Andrew Oxman. GRADE Handbook. https://gdt.gradepro.org/app/handbook/handbook.html#h.f7lc8w9c3nh8 (accessed Feb 8, 2025).

19 Higgins J, Thomas J, Chandler J, *et al.*, editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). 2023. www.training.cochrane.org/handbook.

20 Viechtbauer W. Conducting Meta-Analyses in *R* with the **metafor** Package. *J Stat Softw* 2010; **36**. DOI:10.18637/jss.v036.i03.

21 R Core Team. R: A language and environment for statistical computing. 2023. https://www.R-project.org/.

22 Whiting PF. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011; **155**: 529.

23 Adeli K, Higgins V, Nieuwesteeg M, *et al.* Complex Reference Values for Endocrine and Special Chemistry Biomarkers across Pediatric, Adult, and Geriatric Ages: Establishment of Robust Pediatric and Adult Reference Intervals on the Basis of the Canadian Health Measures Survey. *Clin Chem* 2015; **61**: 1063–74.

24 Bailey D, Colantonio D, Kyriakopoulou L, *et al.* Marked Biological Variance in Endocrine and Biochemical Markers in Childhood: Establishment of Pediatric Reference Intervals Using Healthy Community Children from the CALIPER Cohort. *Clin Chem* 2013; **59**: 1393–405.

25 Bogner B, Schwenoha K, Vogl M, *et al.* Evaluation of reference intervals of haematological and biochemical markers in an Austrian adolescent study cohort. *Clin Chem Lab Med CCLM* 2019; **57**: 891–900.

26 Bohn MK, Higgins V, Asgari S, *et al.* Paediatric reference intervals for 17 Roche cobas 8000 e602 immunoassays in the CALIPER cohort of healthy children and adolescents. *Clin Chem Lab Med CCLM* 2019; **57**: 1968–79.

27 Bohn MK, Horn P, League D, Steele P, Hall A, Adeli K. Pediatric reference intervals for endocrine markers and fertility hormones in healthy children and adolescents on the Siemens Healthineers Atellica immunoassay system. *Clin Chem Lab Med CCLM* 2021; **59**: 1421–30.

28 Chinn S, Rona R, Duran-Tauleria E. Serum cholesterol and haematology at age eight to ten years. *Scand J Clin Lab Invest* 1998; **58**: 135–42.

29 Elmlinger MW, Kühnel W, Ranke MB. Reference Ranges for Serum Concentrations of Lutropin (LH), Follitropin (FSH), Estradiol (E2), Prolactin, Progesterone, Sex Hormone-Binding Globulin (SHBG), Dehydroepiandrosterone Sulfate (DHEAS), Cortisol and Ferritin in Neonates, Children and Young Adults. *Clin Chem Lab Med* 2002; **40**. DOI:10.1515/cclm.2002.202.

30 Hall A, Bohn MK, Wilson S, Higgins V, Adeli K. Continuous reference intervals for 19 endocrine, fertility, and immunochemical markers in the CALIPER cohort of healthy children and adolescents. *Clin Biochem* 2021; **94**: 35–41.

31 Higgins V, Fung AWS, Chan MK, Macri J, Adeli K. Pediatric reference intervals for 29 Ortho VITROS 5600 immunoassays using the CALIPER cohort of healthy children and adolescents. *Clin Chem Lab Med CCLM* 2018; **56**: 327–40.

32 Hollowell JG, van Assendelft OW, Gunter EW, *et al.* Hematological and iron-related analytes--reference data for persons aged 1 year and over: United States, 1988-94. *Vital Health Stat 11* 2005; : 1–156.

33 Hoq M, Matthews S, Karlaftis V, *et al.* Reference Values for 30 Common Biochemistry Analytes Across 5 Different Analyzers in Neonates and Children 30 Days to 18 Years of Age. *Clin Chem* 2019; **65**: 1317–26.

34 Karbasy K, Lin DCC, Stoianov A, *et al.* Pediatric reference value distributions and covariate-stratified reference intervals for 29 endocrine and special chemistry biomarkers on the Beckman Coulter Immunoassay Systems: a CALIPER study of healthy community children. *Clin Chem Lab Med CCLM* 2016; **54**. DOI:10.1515/cclm-2015-0558.

35 Larsson SM, Hillarp A, Hellström-Westas L, Domellöf M, Lundahl T, Andersson O. When age really matters; ferritin reference intervals during infancy revisited. *Scand J Clin Lab Invest* 2019; **79**: 590–4.

36 Parkin PC, Hamid J, Borkhoff CM, *et al.* Laboratory reference intervals in the assessment of iron status in young children. *BMJ Paediatr Open* 2017; **1**: e000074.

37 Rieger K, Vogel M, Engel C, *et al.* Reference intervals for iron-related blood parameters: results from a population-based cohort study (LIFE Child). *LaboratoriumsMedizin* 2016; **40**: 000010151520160019.

38 Southcott EK, Kerrigan JL, Potter JM, *et al.* Establishment of pediatric reference intervals on a large cohort of healthy children. *Clin Chim Acta* 2010; **411**: 1421–7.

39 Tahmasebi H, Asgari S, Hall A, *et al.* Influence of ethnicity on biochemical markers of health and disease in the CALIPER cohort of healthy children and adolescents. *Clin Chem Lab Med CCLM* 2020; **58**: 605–17.

40 Valberg LS, Sorbie J, Ludwig J, Pelletier O. Serum ferritin and the iron status of Canadians. *Can Med Assoc J* 1976; **114**: 417–21.

41 Wiedemann G, Jonetz-Mentzel L. Establishment of reference ranges for ferritin in neonates, infants, children and adolescents. *Eur J Clin Chem Clin Biochem J Forum Eur Clin Chem Soc* 1993; **31**: 453–7.

42 Wilson SM, Bohn MK, Madsen A, Hundhausen T, Adeli K. LMS-based continuous reference percentiles for 14 laboratory parameters in the CALIPER cohort of healthy children and adolescents. *Clin Chem Lab Med CCLM* 2023; **61**: 1105–15.

43 Abdullah K, Birken CS, Maguire JL, *et al.* Re-Evaluation of Serum Ferritin Cut-Off Values for the Diagnosis of Iron Deficiency in Children Aged 12-36 Months. *J Pediatr* 2017; **188**: 287–90.

44 Mei Z, Addo OY, Jefferds ME, Sharma AJ, Flores-Ayala RC, Brittenham GM. Physiologically based serum ferritin thresholds for iron deficiency in children and non-pregnant women: a US National Health and Nutrition Examination Surveys (NHANES) serial cross-sectional study. *Lancet Haematol* 2021; **8**: e572–82.

45 Mukhtarova N, Ha B, Diamond CA, Plumb AJ, Kling PJ. Serum Ferritin Threshold for Iron Deficiency Screening in One-Year-Old Children. *J Pediatr* 2022; **245**: 217–21.

46 Parkin PC, Koroshegyi C, Mamak E, *et al.* Association between Serum Ferritin and Cognitive Function in Early Childhood. *J Pediatr* 2020; **217**: 189-191.e2.

47 Truong J, Naveed K, Beriault D, Lightfoot D, Fralick M, Sholzberg M. The origin of ferritin reference intervals: a systematic review. *Lancet Haematol* 2024; **11**: e530–9.

48 Wozniak, M., Catomeris, P., Li, H., *et al.* Guidelines for the Use of Laboratory Tests for Iron Deficiency (CLP 002). Ontario Association of Medical Laboratories, 2024.

Supplemental tables and figures

Table S1. Characteristics of the included studies

Author	Year	Country	No. of sample	Analyzer	Ref. Type	Age Range	90%CI	CLSI Cited	Daly 2017 Cited	CLSI Followed
Bohn	2023	Canada	670	Mindray	RI	1 m - 19 y	Yes	Yes	No	Yes
Li	2023	China	497	Mindray	RI	0 m - 1 m	Yes	Yes	Yes	Yes
Song	2022	China	11750	Sysmex	RI	1 m - 18 y	Yes	Yes	Yes	Yes
Yan	2022	China	12823	Sysmex	RC	3 m - 17 y	Yes	Yes	Yes	Yes
Baudouin	2021	Côte d'Ivoire	310	Sysmex	RI	5 y - 12 y	Yes	No	No	No
Hamid	2021	Canada	2012	Sysmex	RI	1 m - 36 m	Yes	Yes	Yes	Yes
Ouma	2021	Kenya	1509	Beckman Coulter	RI	1 m - 17 m	Yes	Yes	No	No
Ringoringo	2021	Indonesia	277	Sysmex	RI	0 m - 0 m	No	No	No	No
Sachdev	2021	India	38920	Not provided	RC	12 m - 18 y	No	No	No	No
Wilson	2021	Canada	496	Sysmex	RC	6 m - 18 y	No	Yes	No	Yes
Beavogui	2020	Guinea	802	Horiba	RI	6 y - 16 y	No	Yes	No	No
Bohn	2020	Canada	634	Sysmex	RI	0 m - 21 y	Yes	Yes	No	Yes
Haileslasie	2020	Ethiopia	249	Sysmex	RI	12 y - 18 y	Yes	Yes	No	Yes
Kieh	2020	Liberia	694	Abbott	RI	6 y - 18 y	No	Yes	No	No
Sun	2020	China	9944	Sysmex	RI	6 m - 7 y	Yes	Yes	No	Yes
Bogner	2019	Austria	102	Sysmex	RI	14 y - 18 y	Yes	Yes	No	No
Fulgoni	2019	US	21285	Beckman Coulter	RI	12 m - 18 y	Yes, No	Yes	No	Yes
Bimerew	2018	Ethiopia	334	Sysmex	RI	5 y - 18 y	No	Yes	No	No
Yanamandra	2018	India	335	Sysmex	RI	4 y - 20 y	Yes	No	No	No
Parkin	2017	Canada	2600	Sysmex	RI	3 y - 11 y	Yes	Yes	No	Yes
Wang	2017	China	616	Sysmex	RI	4 m - 7 y	Yes	Yes	No	Yes
Du	2016	China	698	Sysmex	RI	7 y - 14 y	Yes	No	No	No
Imoru	2016	Nigeria	72	Other	RI	0 m - 0 m	Yes	No	NA	No
Lee	2016	South Korea	534	Sysmex	RI	12 m - 24 m	Yes	Yes	No	Yes
Rieger	2016	Germany	1779	Sysmex	RC	4 y - 17 y	No	Yes	No	No
Adeli_a	2015	Canada	2570	Sysmex	RI	3 y - 20 y	Yes	Yes	No	Yes
Gomani	2015	Zimbabwe	224	Sysmex	RI	12 y - 18 y	Yes	Yes	No	Yes
Lim	2015	South Korea	1865	Sysmex	RI	10 y - 18 y	Yes	No	No	No
Odhiambo	2015	Kenya	242	Beckman Coulter	RI	16 y - 18 y	No	Yes	No	No
Dosoo	2014	Ghana	1272	Horiba	RI	6 m - 18 y	No	Yes	No	No
Odutola	2014	Gambia	696	Abbott	RI	2 m - 10 m	No	No	No	No
Aldrimer	2013	Sweden	253	Abbott	RI	12 y - 19 y	No	Yes	No	No
Melioli	2011	Italy	820	Beckman Coulter	RI	0 m - 0 m	No	No	No	No
Zeh	2011	Kenya	133	Beckman Coulter	RI	13 y - 18 y	No	Yes	No	No
Buchanan	2010	Tanzania	600	Beckman Coulter	RI	0 m - 18 y	No	Yes	No	No
Buseri	2010	Nigeria	1021	Not provided	RI	0 m - 18 y	Yes	No	NA	No

Author	Year	Country	No. of sample	Analyzer	Ref. Type	Age Range	90%CI	CLSI Cited	Daly 2017 Cited	CLSI Followed
Hollowell	2005	US	10245	Beckman Coulter	RI	12 m - 20 y	Yes	No	No	No
El-Hazmi	2001	Saudi Arabia	1526	Beckman Coulter	RI	0 m - 15 y	Yes	No	No	No
Flegar-Mestric	2000	Croatia	996	Beckman Coulter	RI	8 y - 20 y	No	Yes	No	No
Taylor	1997	Ireland	2120	Beckman Coulter	RI	4 y - 20 y	Yes	No	No	No
Dallman	1984	US	2662	Beckman Coulter	RI	12 m - 18 y	No	No	No	No
Scheer	1981	US	1894	Other	RI	6 y - 15 y	No	No	No	No
Kelly	1977	Canada	601	Beckman Coulter	RI	10 y - 20 y	Yes	No	NA	No
Valberg	1976	Canada	408	Not provided	RI	12 m - 20 y	Yes	No	No	No
Burman	1972	UK	1121	Ortho	RI	3 m - 25 m	Yes	No	NA	No
Natvig	1967	Norway	1203	Not provided	RI	7 y - 21 y	Yes	No	No	No
Hawkins	1954	Canada	1797	Other	RI	6 y - 18 y	Yes	No	NA	No
Færgeman	1938	Sweden	300	Other	RI	8 y - 15 y	Yes	No	No	No

CLSI - Clinical Laboratory Standards Institute; RI – reference interval; RC – reference curve

Table S2. Study adherence to Clinical Laboratory Standards Institute (CLSI) guidelines

First Author Publication Year	Outlier detection	Partitioning	Method for RIs estimation	Estimation of CIs for RIs
Færgemanet al., (1938)	N	Y	Y	Y
Hawkins et al., (1954)	N	Y	N	Y
Natvig et al., (1967)	N	Y	N	Y
Burman et al., (1972)	N	Y	N	Y
Valberg et al., (1976)	N	N	Y	Y
Kelly et al., (1977)	N	Y	Y	Y
Scheer et al., (1981)	N	Y	Y	N
Dallman et al., (1984)	N	Y	Y	Y
Taylor et al., (1997)	Y	Y	N	Y
Flegar-Mestric et al., (2000)	Y	Y	Y	N
El-Hazmi et al., (2001)	N	Y	N	Y
Hollowell et al., (2005)	N	Y	Y	Y
Buseri et al., (2010)	N	N	N	Y
Buchanan et al., (2010)	Y	Y	Y	N
Melioli et al., (2011)	N	Y	Y	N
Zeh et al., (2011)	N	Y	Y	N
Aldrimer et al., (2013)	Y	Y	Y	N
Dosoo et al., (2014)	Y	Y	Y	N
Odutola et al., (2014)	Y	Y	Y	N
Adeli et al., (2015)	Y	Y	Y	Y
Lim et al., (2015)	N	Y	N	Y
Odhiambo et al., (2015)	N	Y	Y	N
Gomani et al., (2015)	Y	Y	Y	Y
Rieger et al., (2016)	NA	NA	NA	N
Imoru et al., (2016)	N	Y	N	Y
Lee et al., (2016)	Y	Y	Y	Y
Du et al., (2016)	N	Y	Y	Y
Wang et al., (2017)	Y	Y	Y	Y
Parkin et al., (2017)	Y	Y	Y	Y
Bimerew et al., (2018)	N	Y	Y	N
Yanamandra et al., (2018)	N	Y	N	Y
Bogner et al., (2019)	Y	Y	N	Y
Fulgoni et al., (2019)	Y	Y	Y	Y
Kieh et al., (2020)	Y	Y	Y	N
Haileslasie et al., (2020)	Y	Y	Y	Y
Bohn et al., (2020)	Y	Y	Y	Y
Beavogui et al., (2020)	Y	Y	Y	N
Sun et al., (2020)	Y	Y	Y	Y

Sachdev et al., (2021)	NA	NA	NA	N
Wilson et al., (2021)	NA	NA	NA	Y
Hamid et al., (2021)	Y	Y	Y	Y
Ringoringo et al., (2021)	N	Y	N	Y
Ouma et al., (2021)	N	Y	Y	Y
Baudouin et al., (2021)	N	Y	N	Y
Yan et al., (2022)	NA	NA	NA	Y
Song et al., (2022)	Y	Y	Y	Y
Bohn et al., (2023)	Y	Y	Y	Y
Li et al., (2023)	Y	Y	Y	Y

RIs – reference intervals; RCs – reference curves

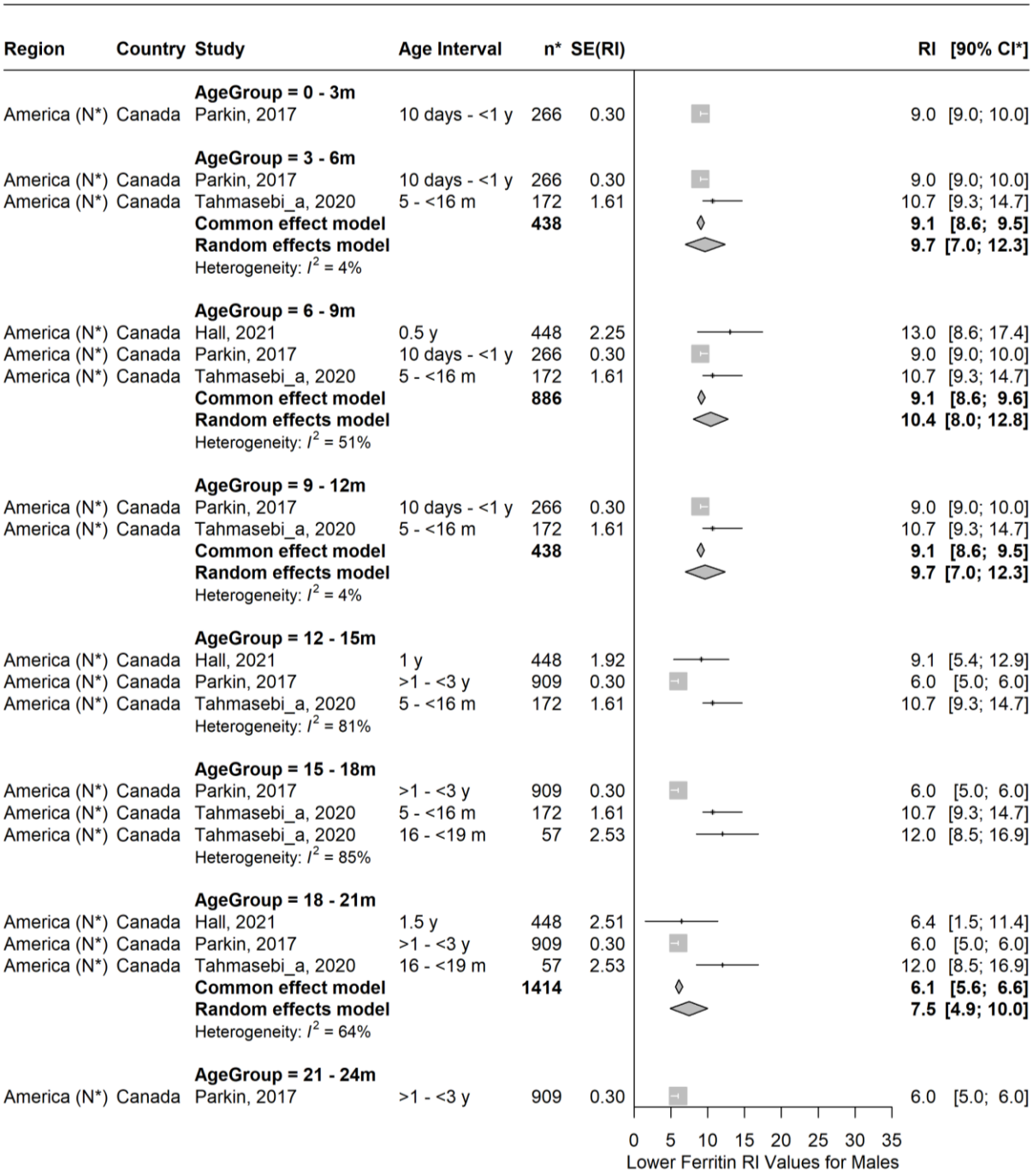


Figure S1. Forest plot of ferritin RIs lower limits for 0-24 months old males

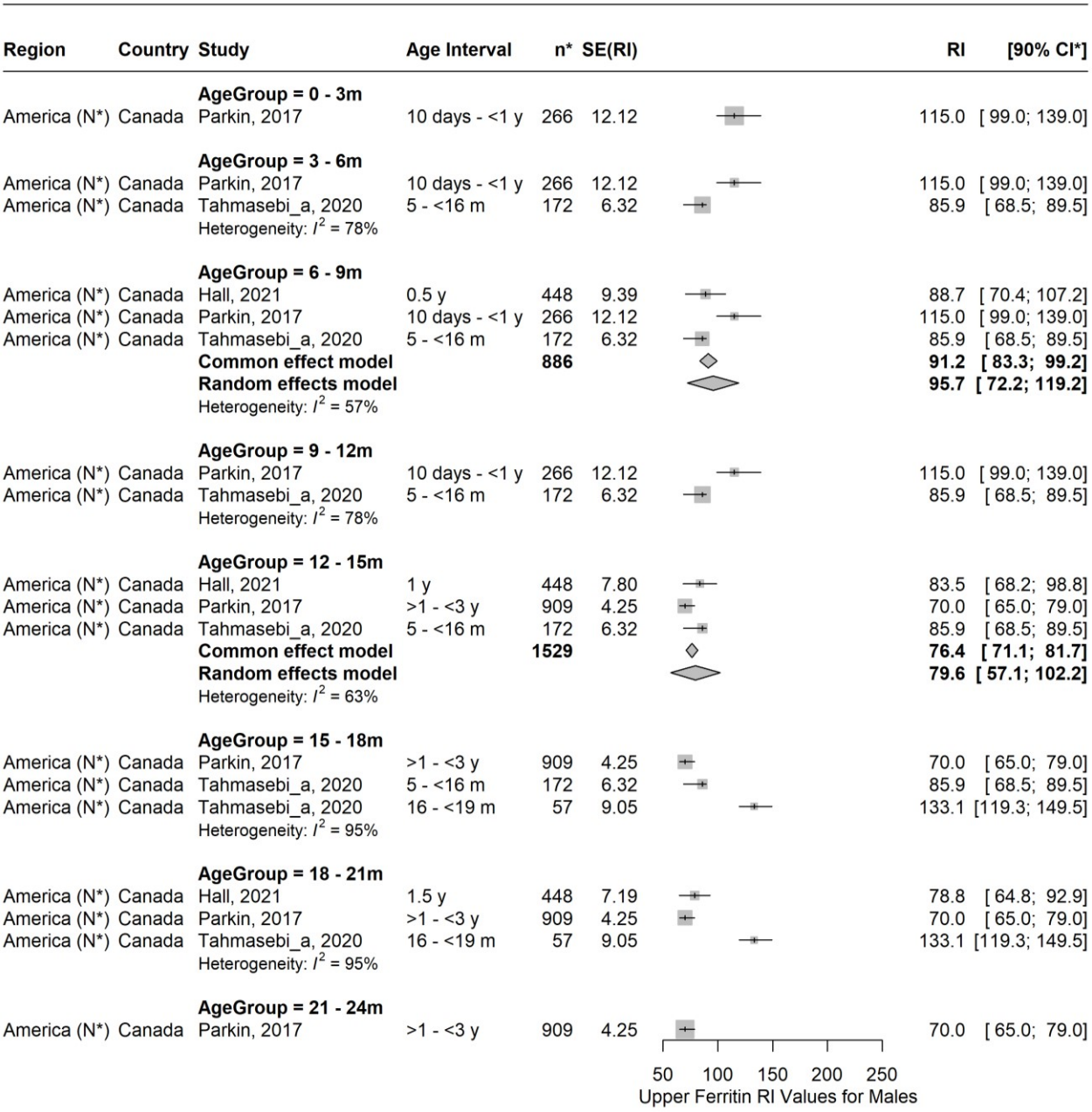


Figure S2. Forest plot of ferritin RIs upper limits for 0-24 months old males

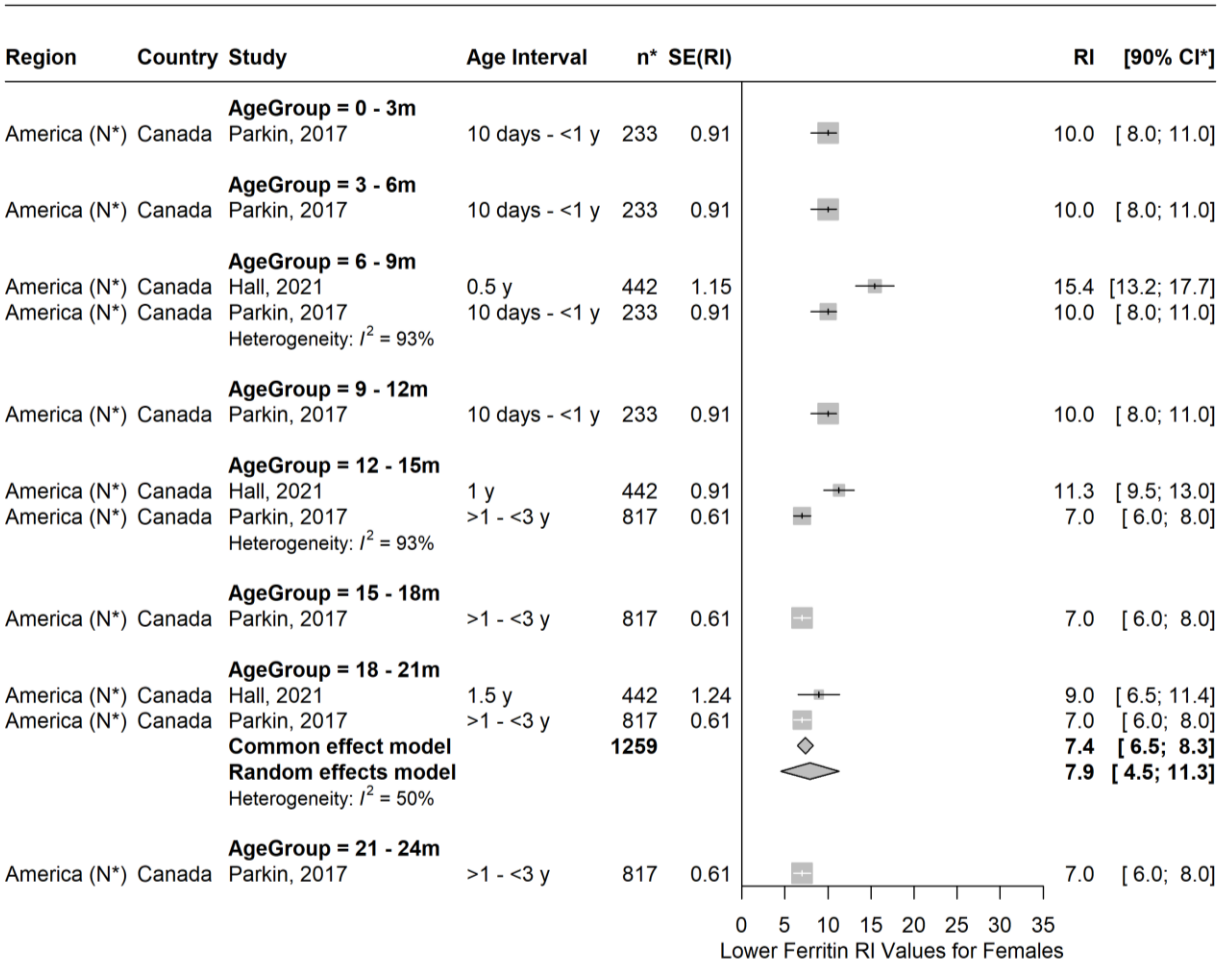


Figure S3. Forest plot of ferritin RIs lower limits for 0-24 months old females

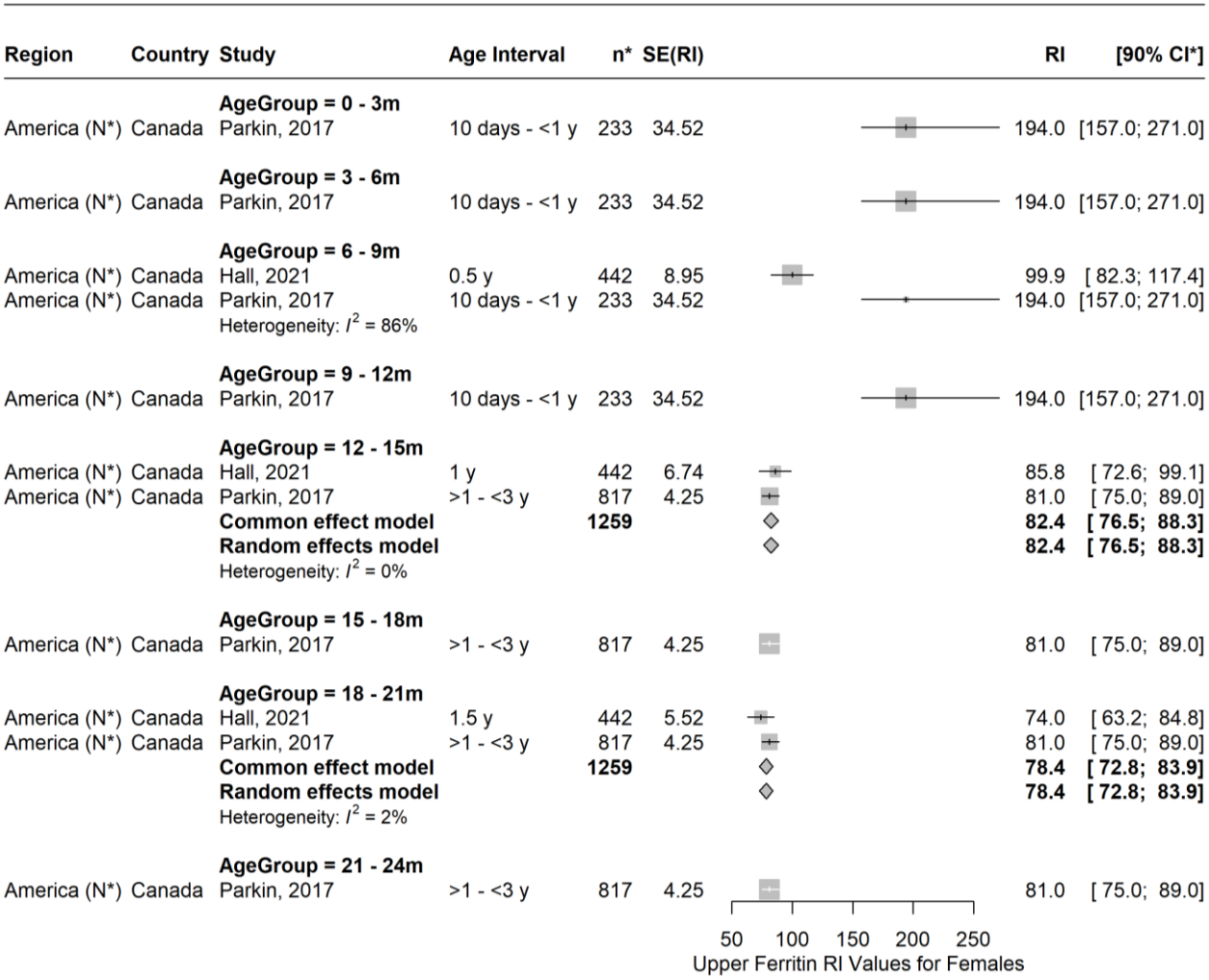


Figure S4. Forest plot of ferritin RIs upper limits for 0-24 months old females

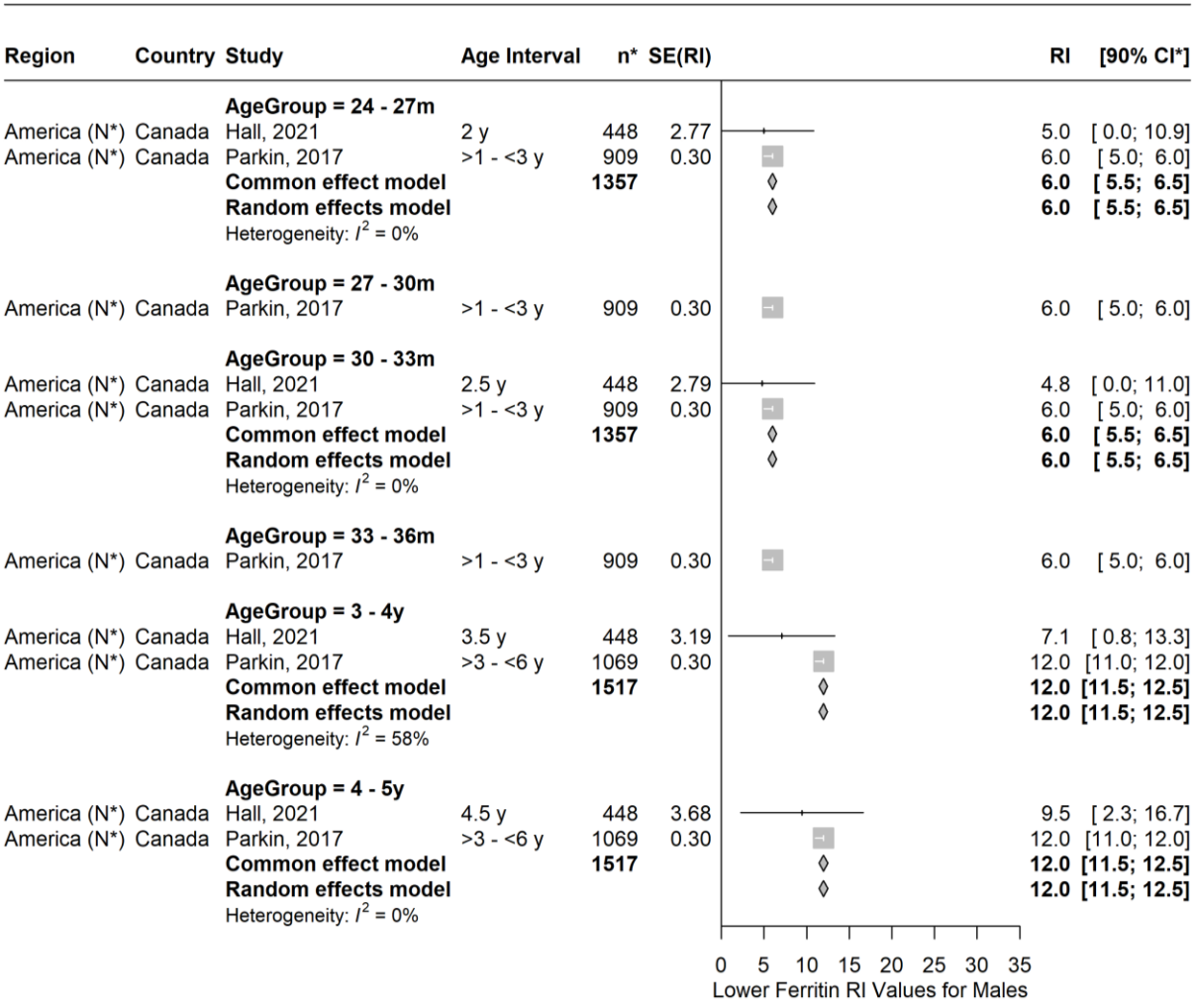


Figure S5. Forest plot of ferritin RIs lower limits for 2-5 years old males

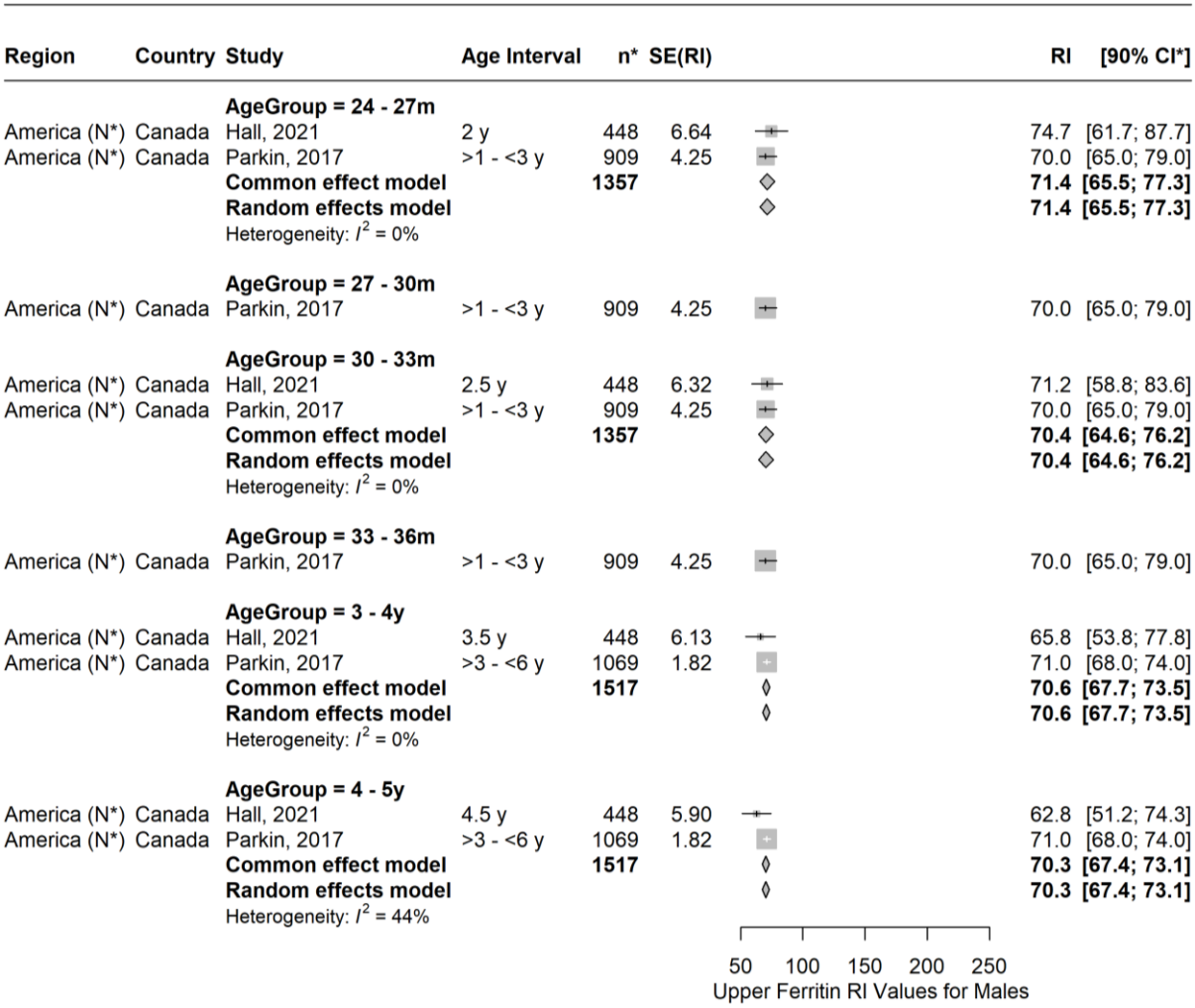


Figure S6. Forest plot of ferritin RIs upper limits for 2-5 years old males

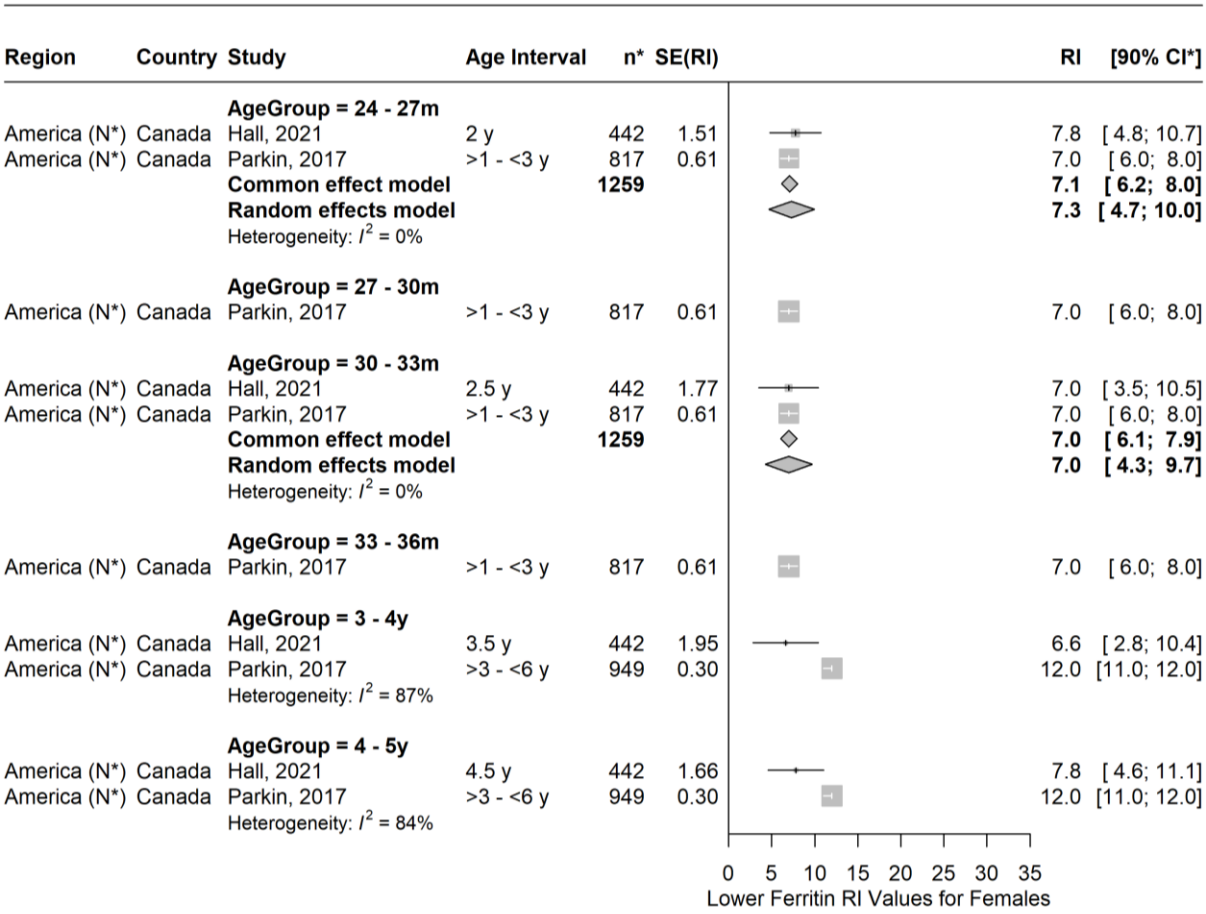


Figure S7. Forest plot of ferritin RIs lower limits for 2-5 years old females

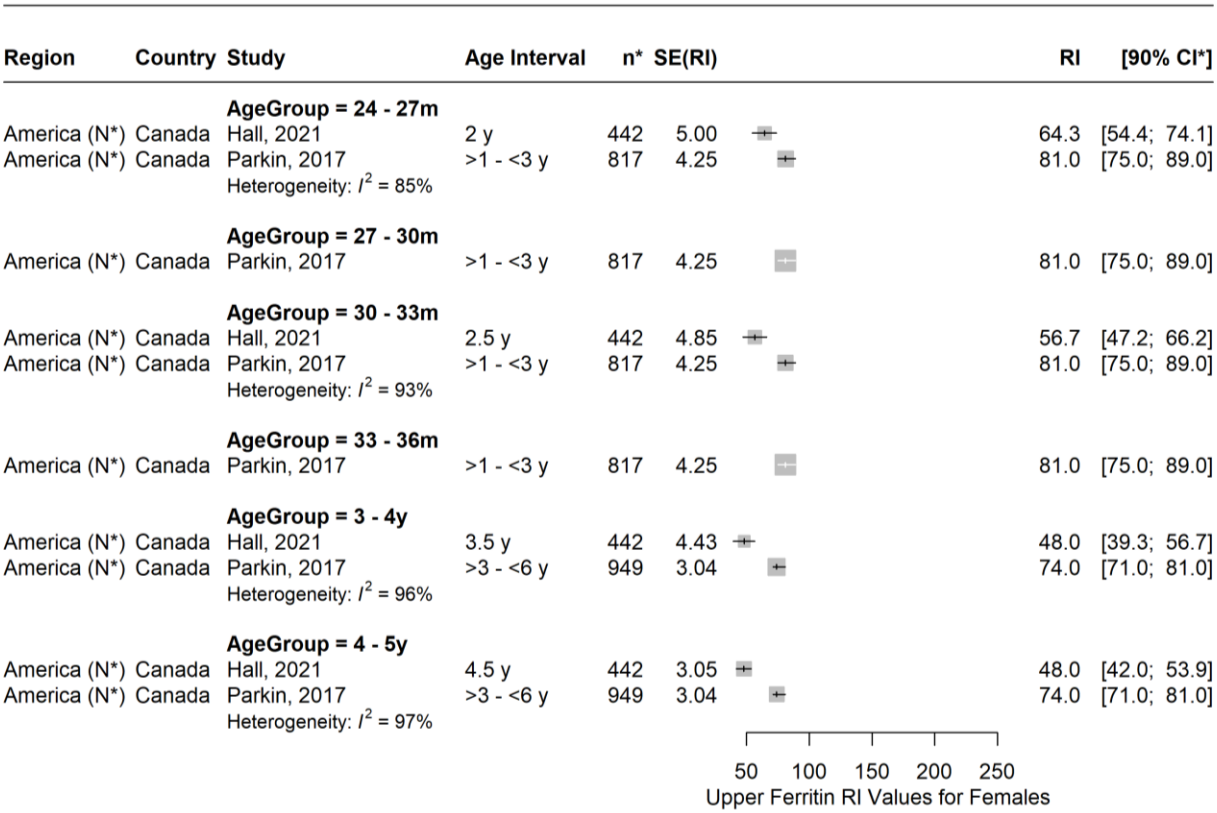


Figure S8. Forest plot of ferritin RIs upper limits for 2-5 years old females

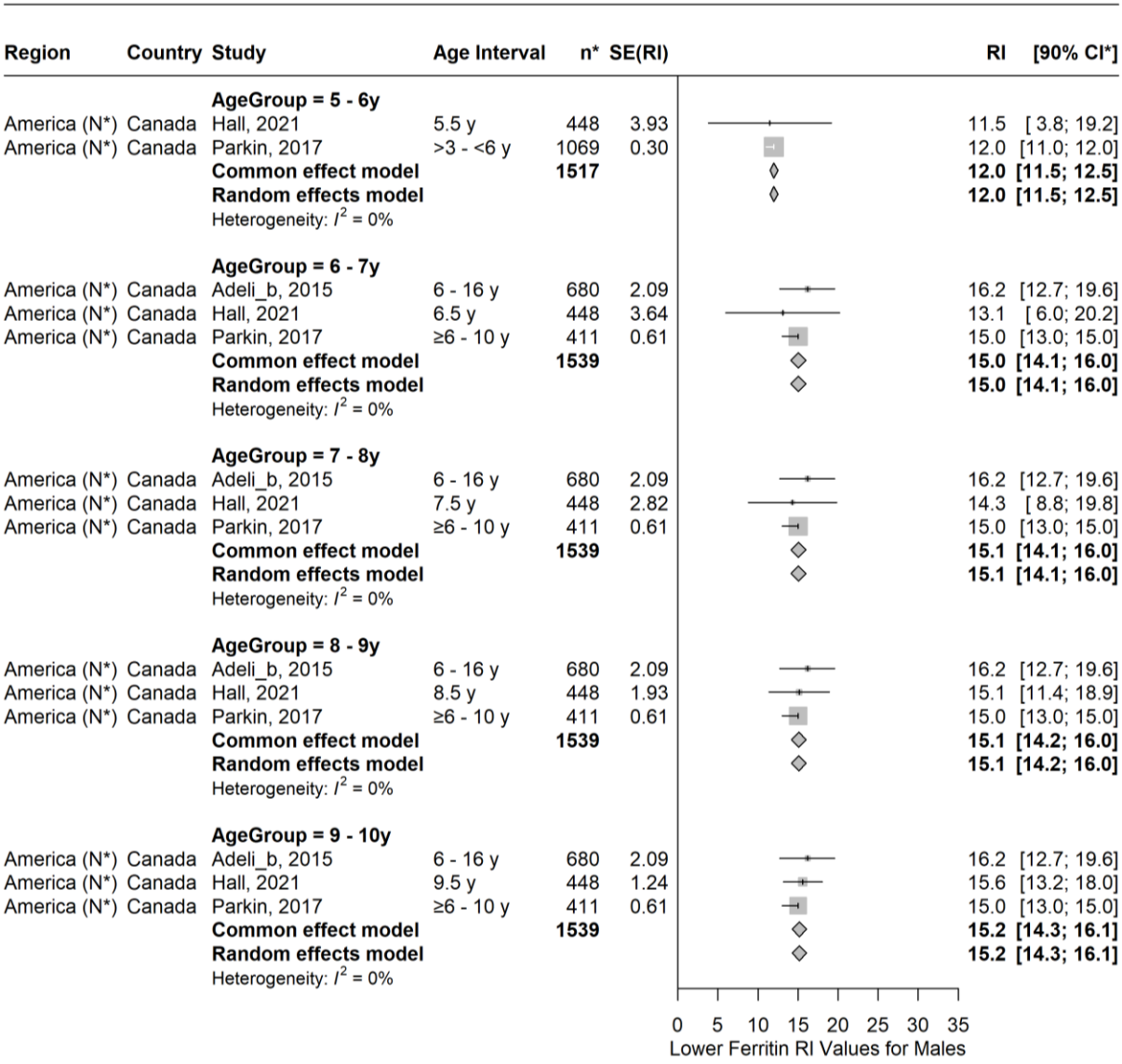


Figure S9. Forest plot of ferritin RIs lower limits for 5-10 years old males

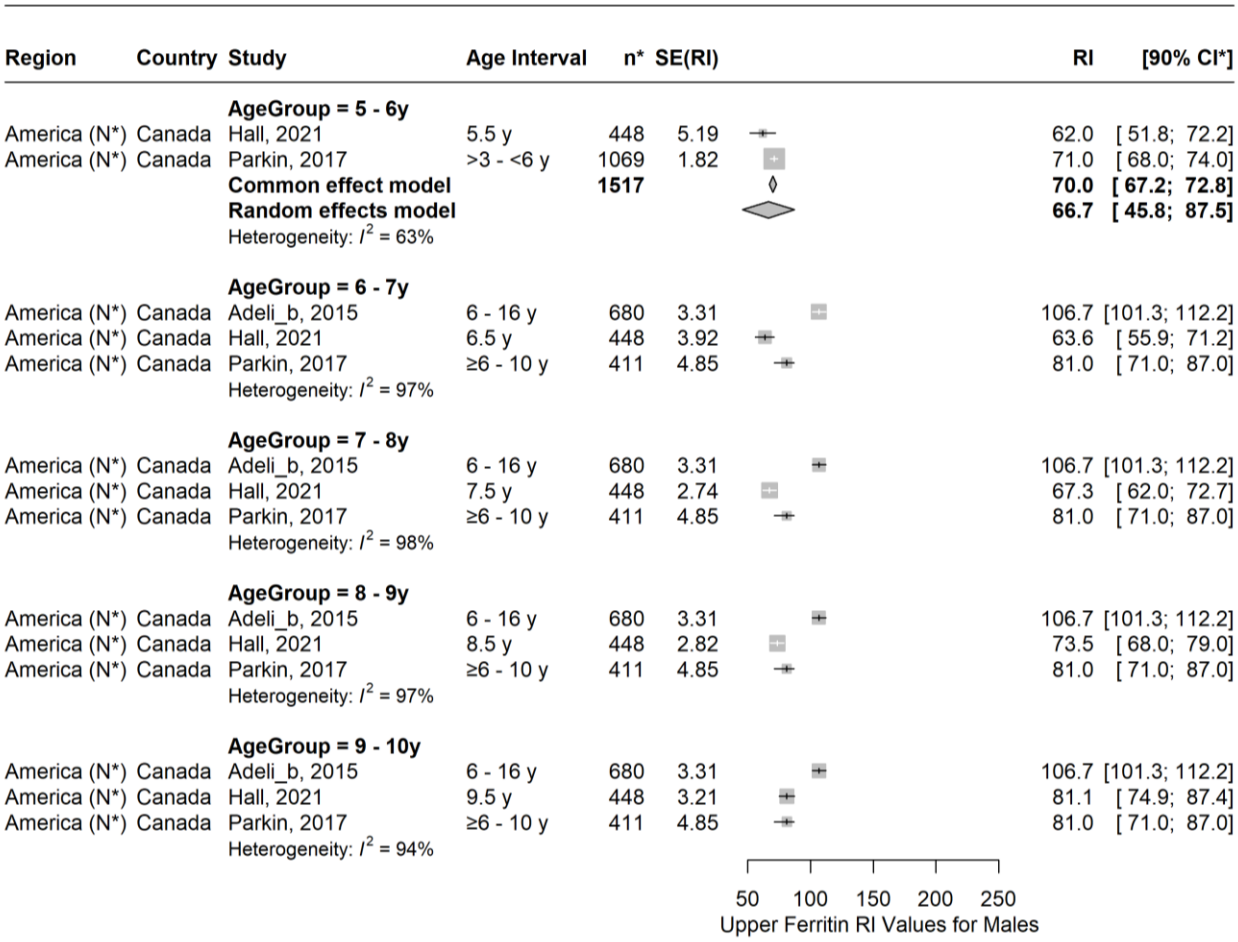


Figure S10. Forest plot of ferritin RIs upper limits for 5-10 years old males

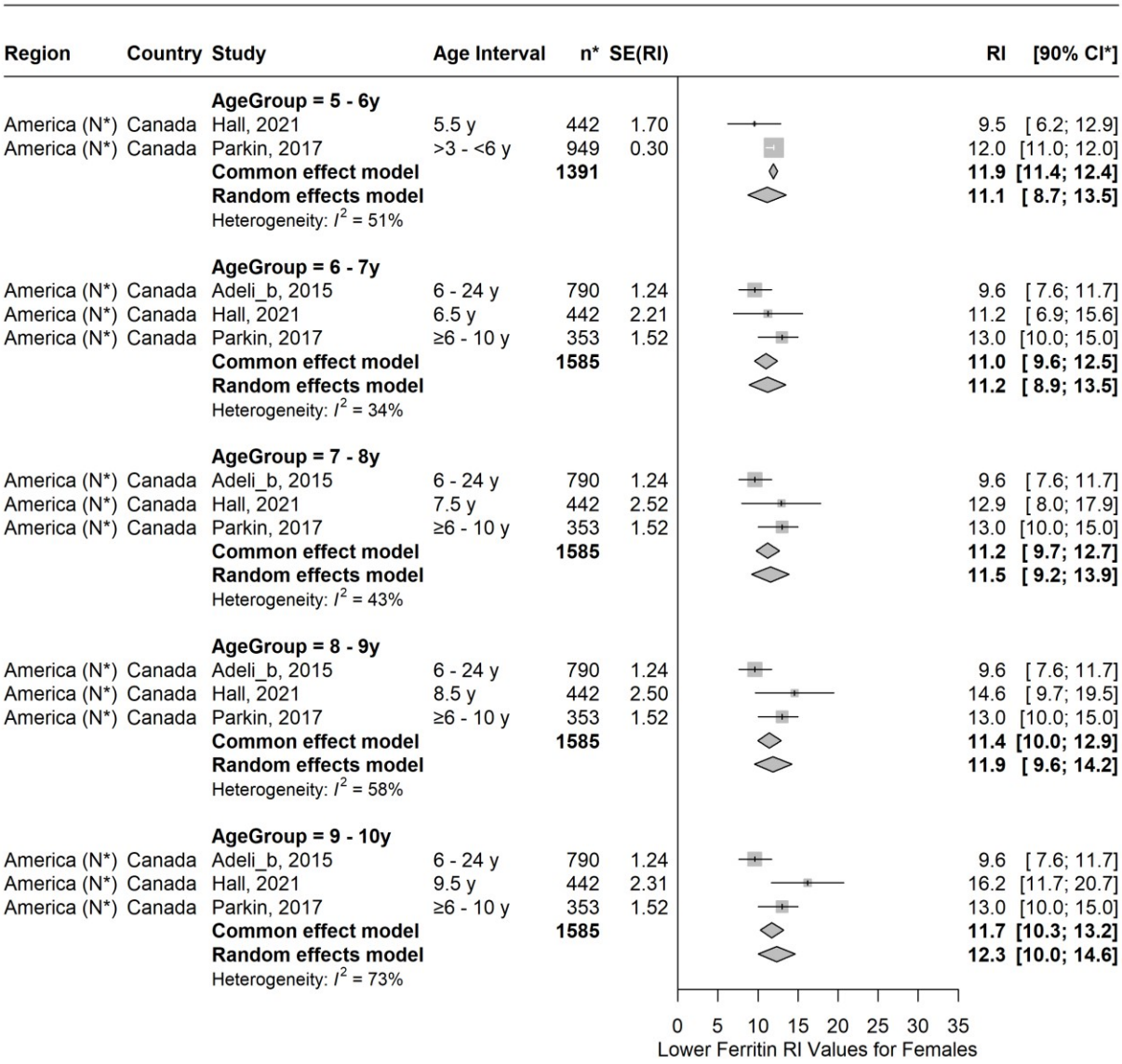


Figure S11. Forest plot of ferritin RIs lower limits for 5-10 years old females

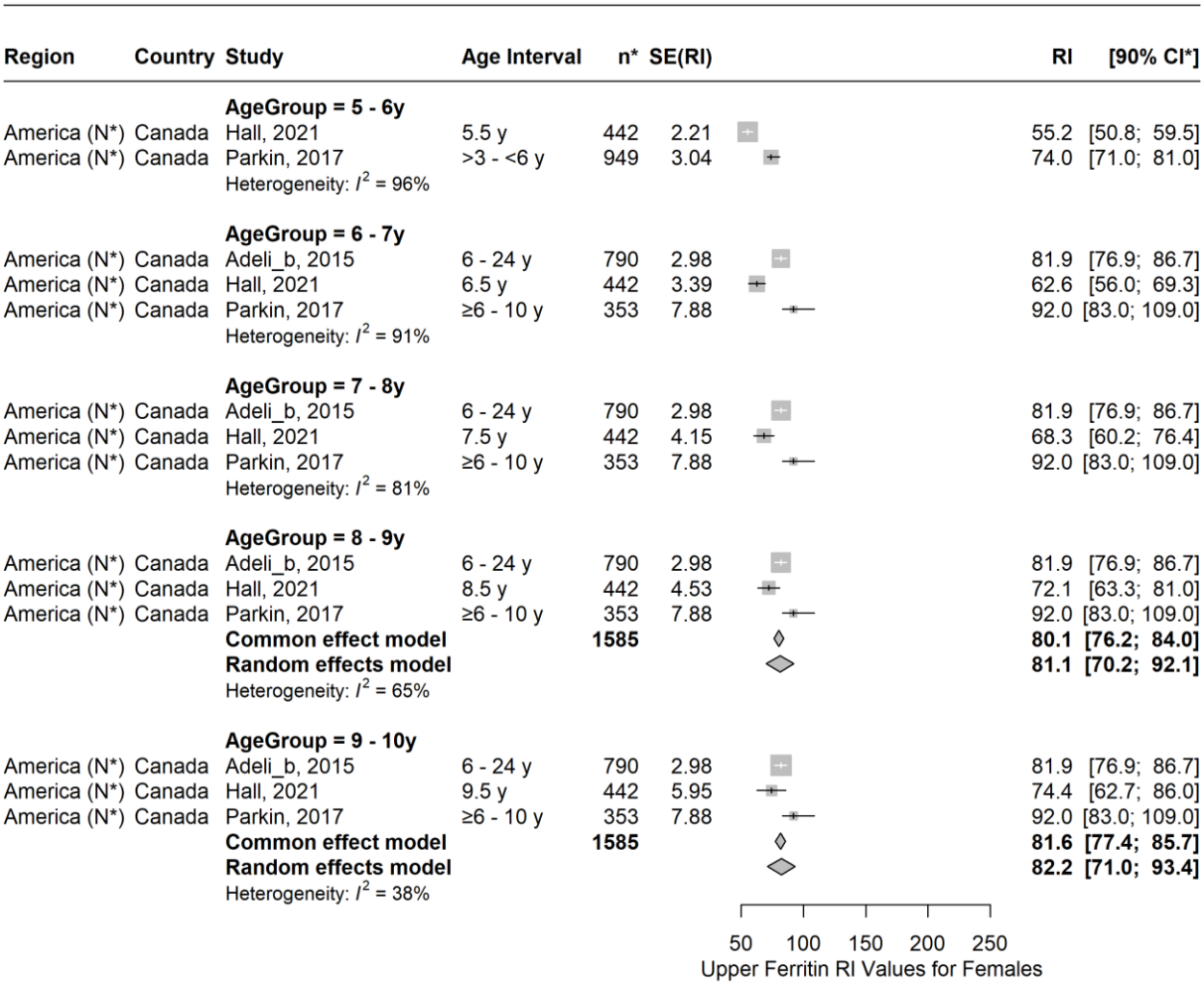


Figure S12. Forest plot of ferritin RIs upper limits for 5-10 years old females

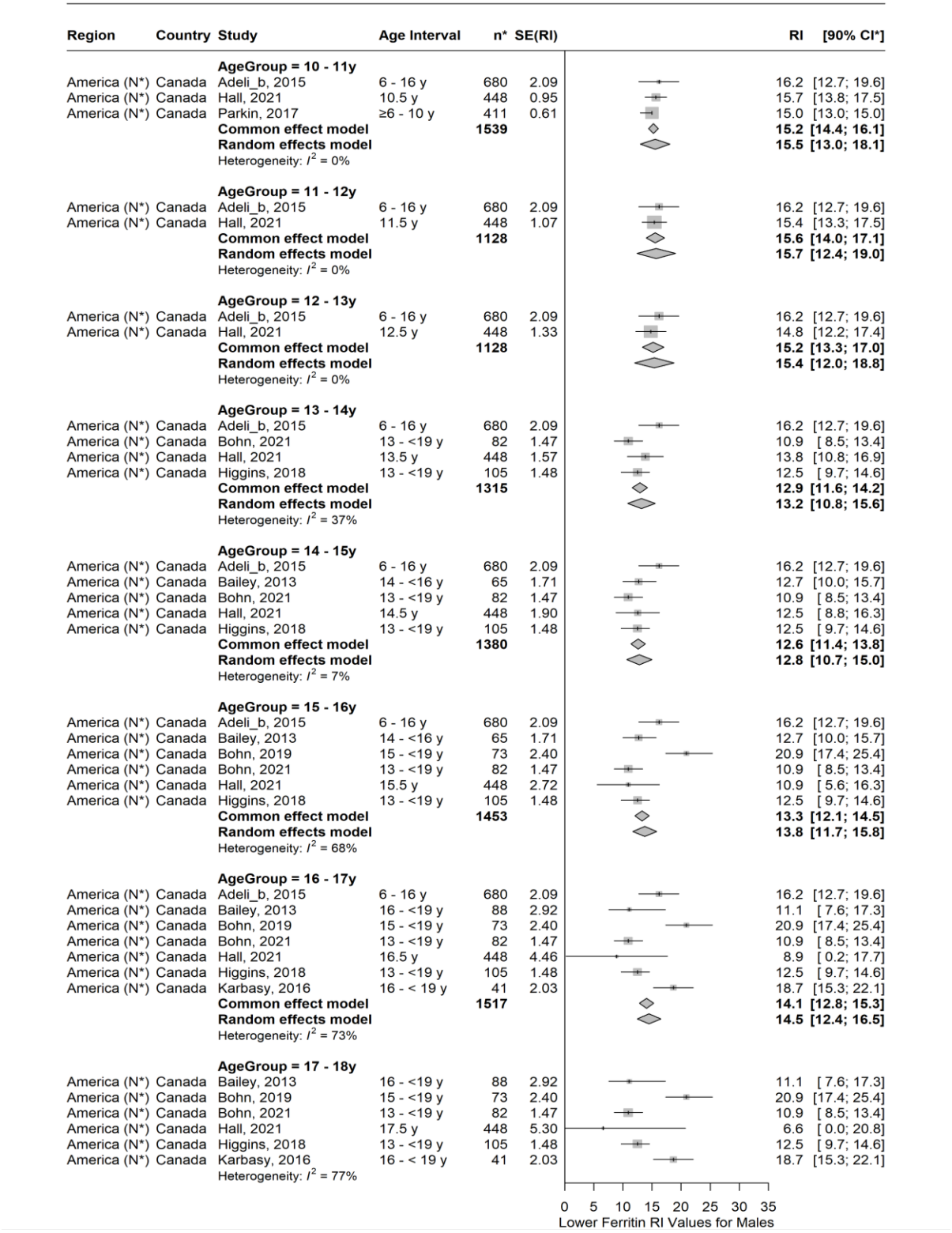


Figure S13. Forest plot of ferritin RIs lower limits for 10-18 years males

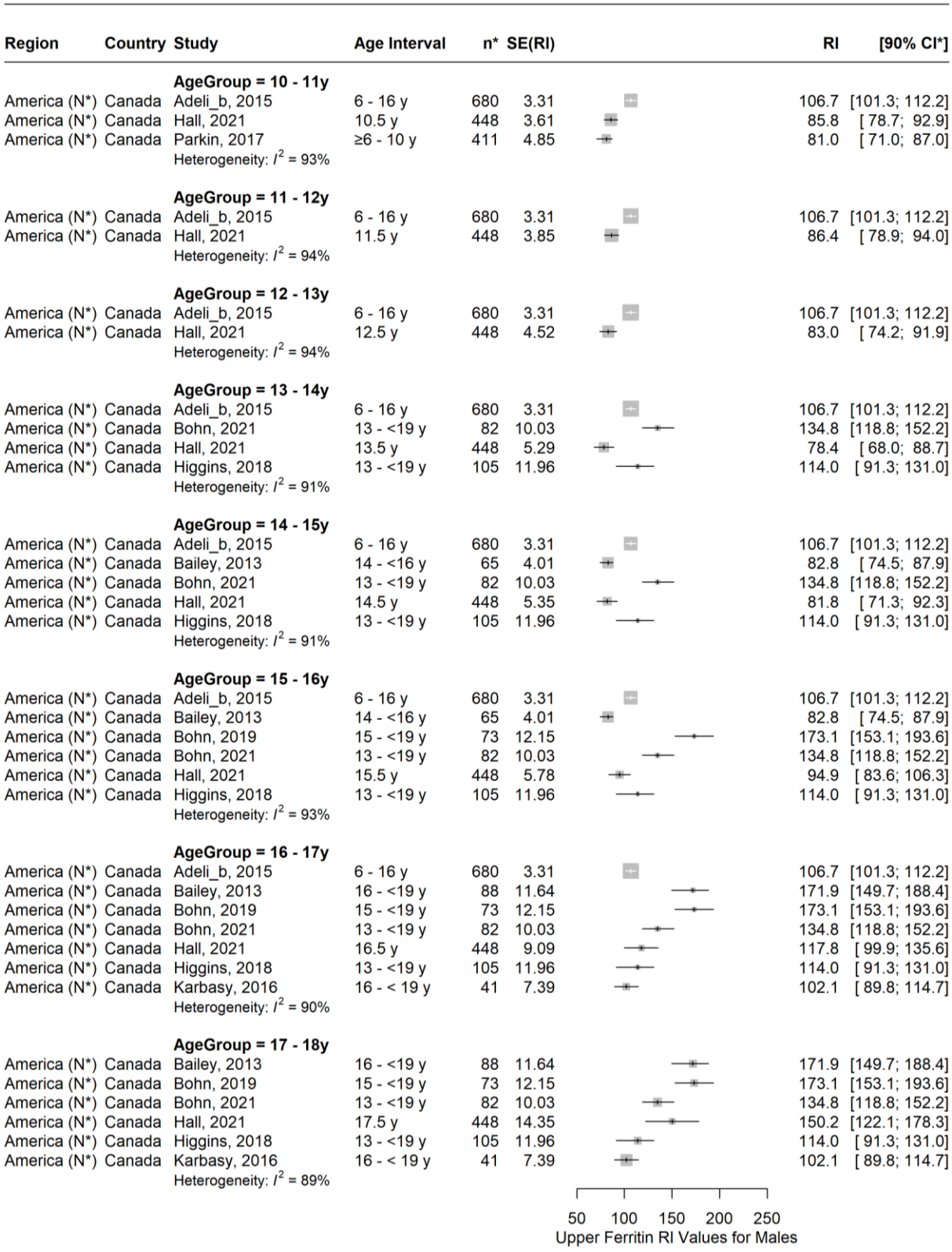


Figure S14. Forest plot of ferritin RIs upper limits for 10-18 years males

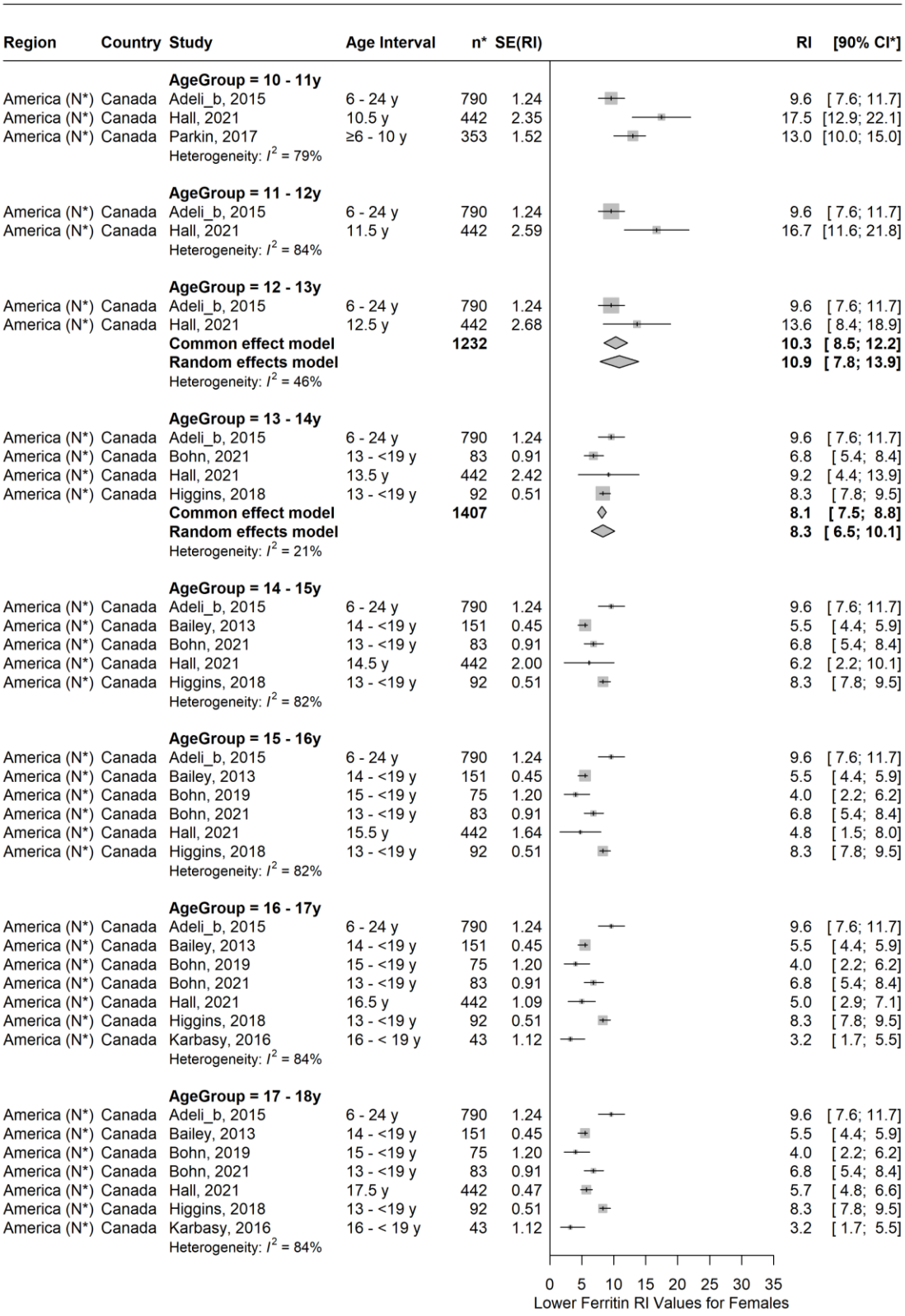


Figure S15. Forest plot of ferritin RIs lower limits for 10-18 years females

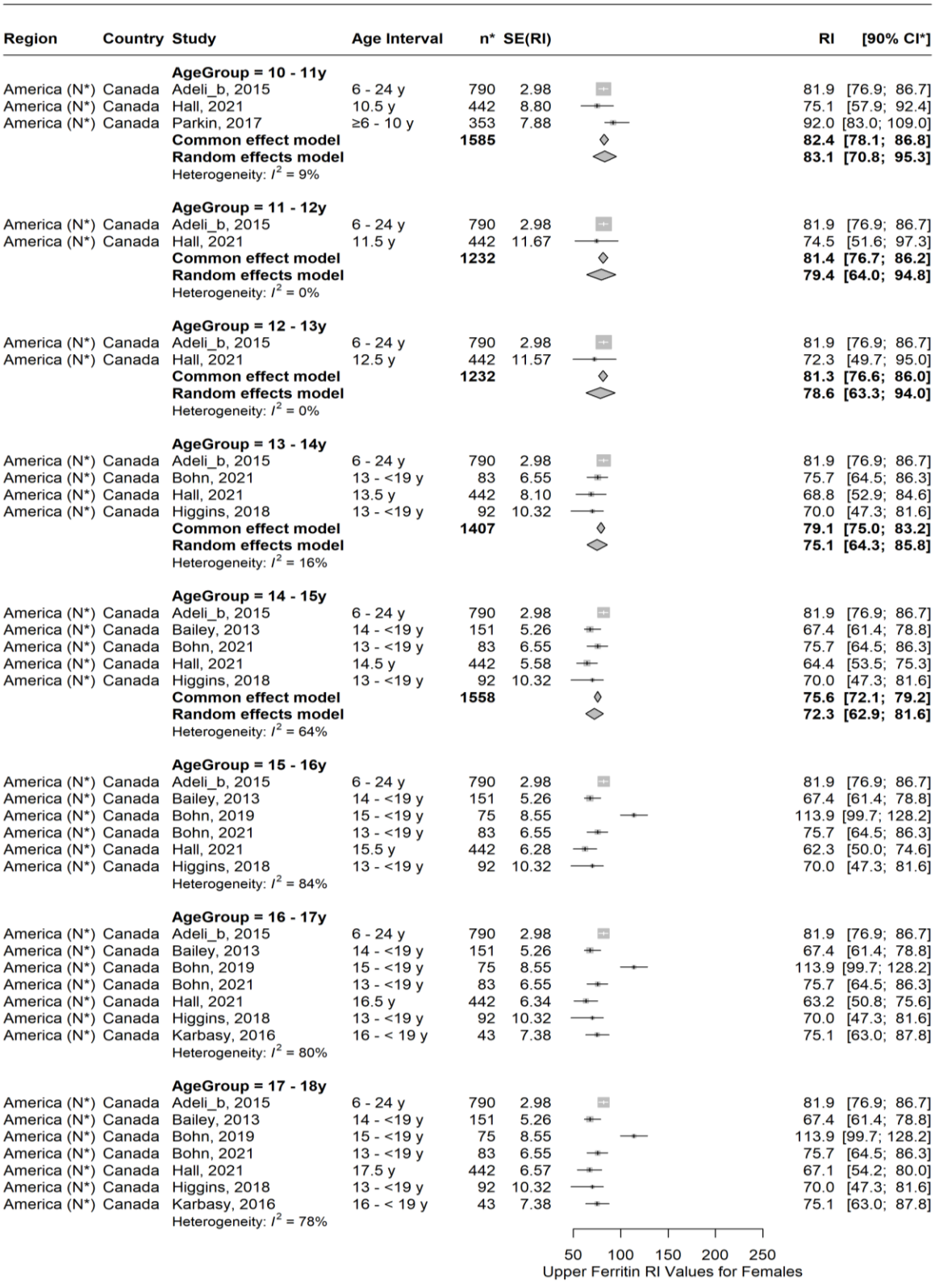


Figure S16. Forest plot of ferritin RIs upper limits for 10-18 years females

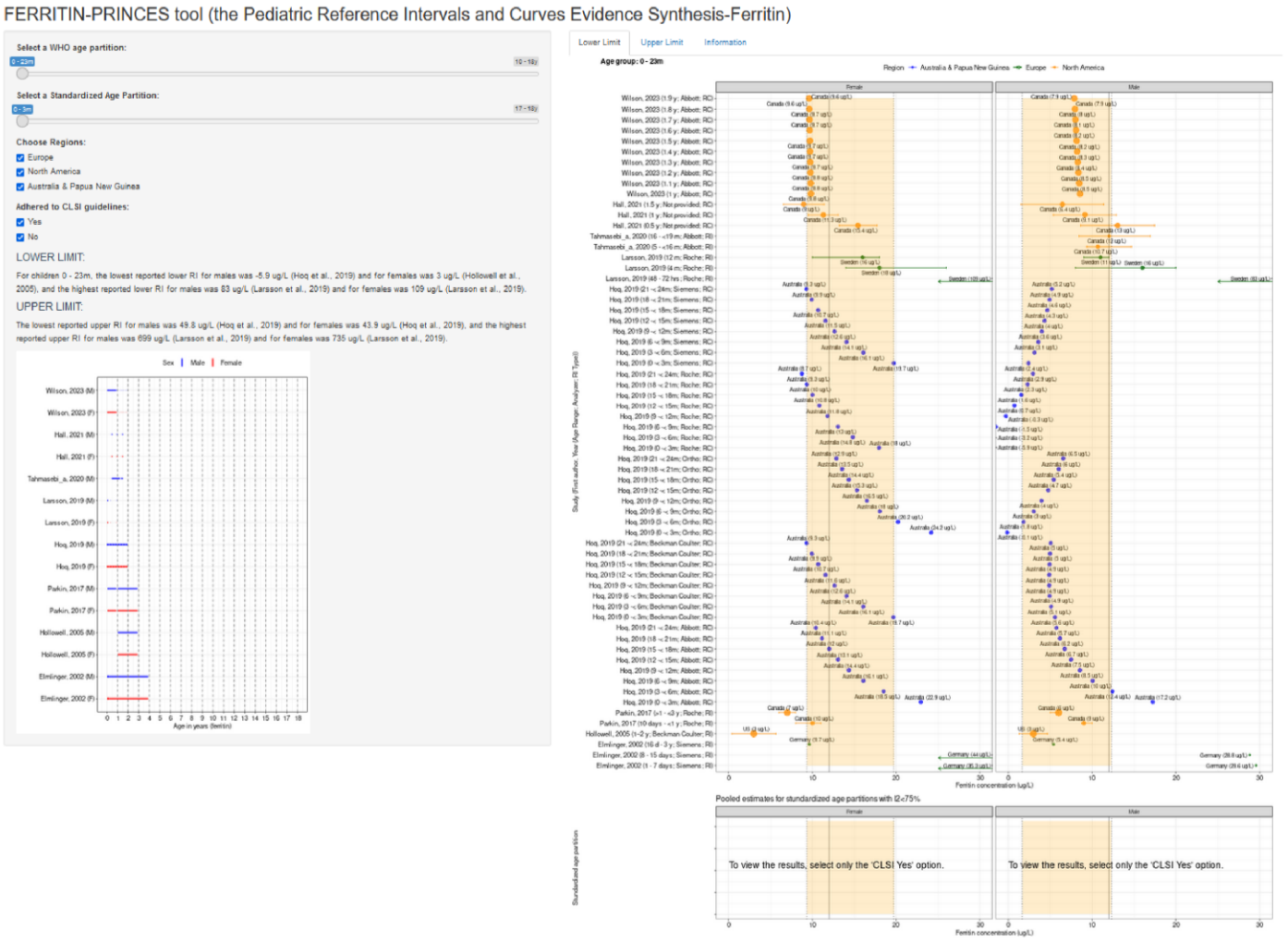


Figure S17. Screenshot of PRINCES-FERRITIN tool

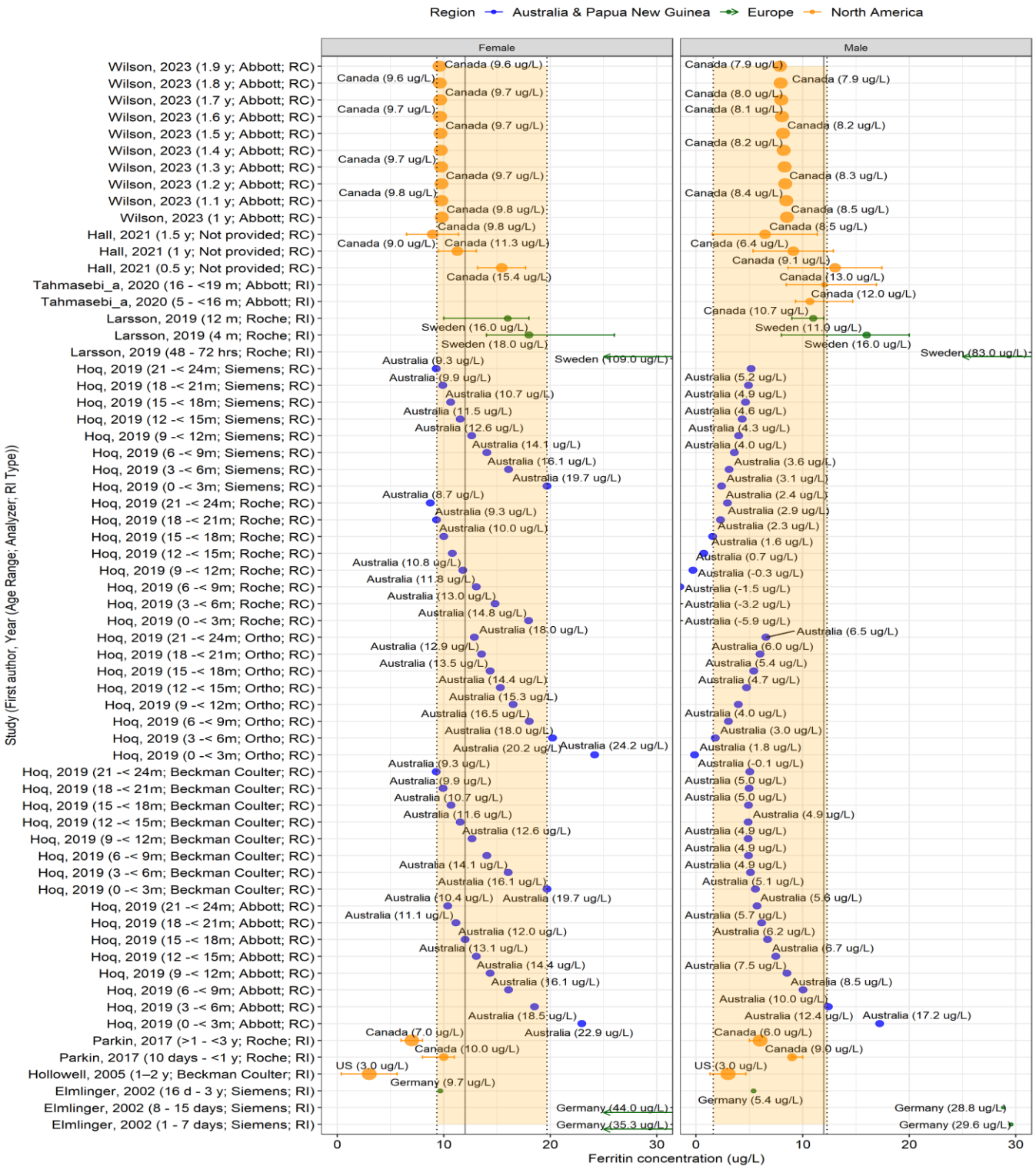


Figure S18. Distribution of ferritin RIs lower limits for 2-5 years old males and females

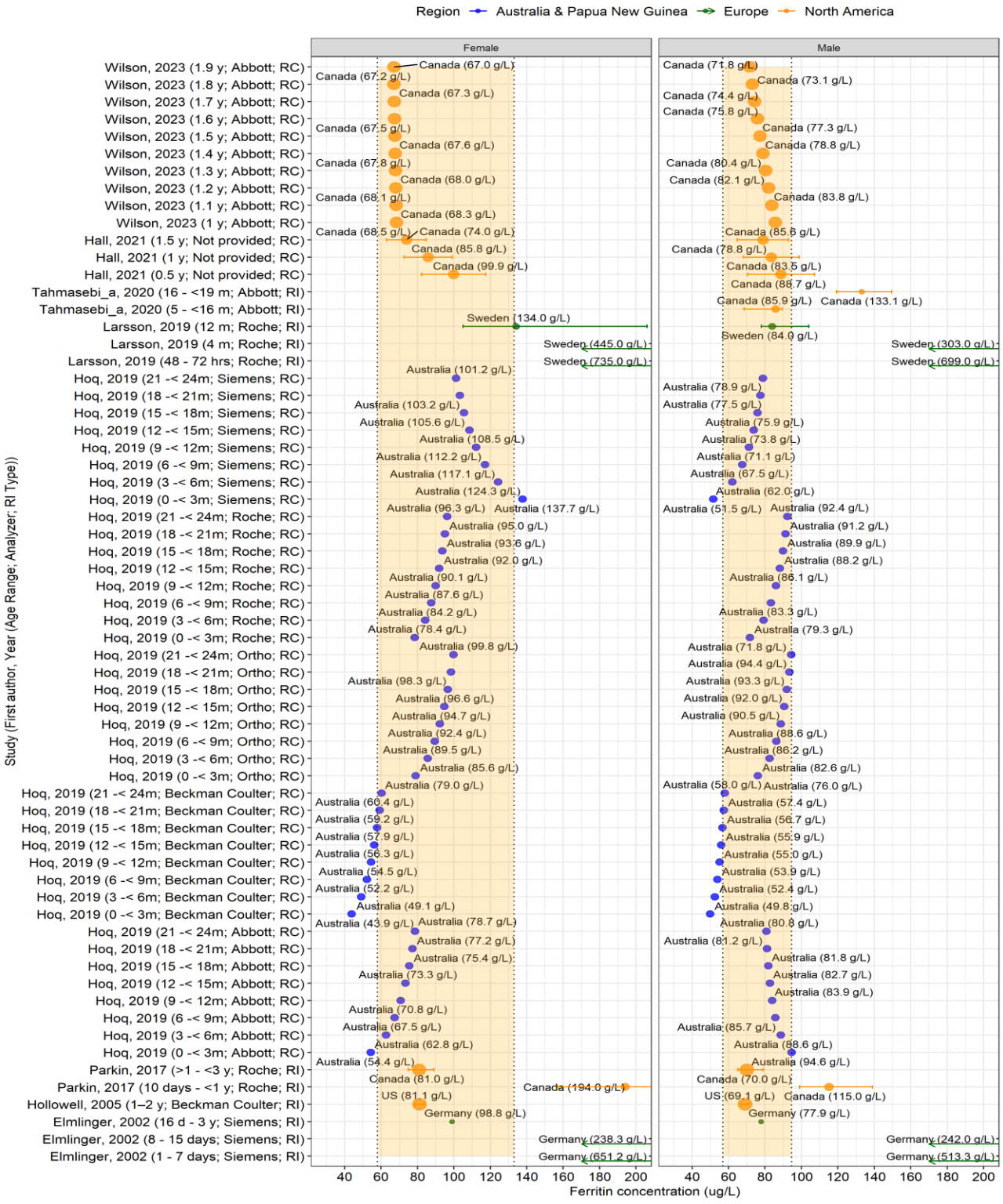


Figure S19. Distribution of ferritin RIs upper limits for 2-5 years old males and females

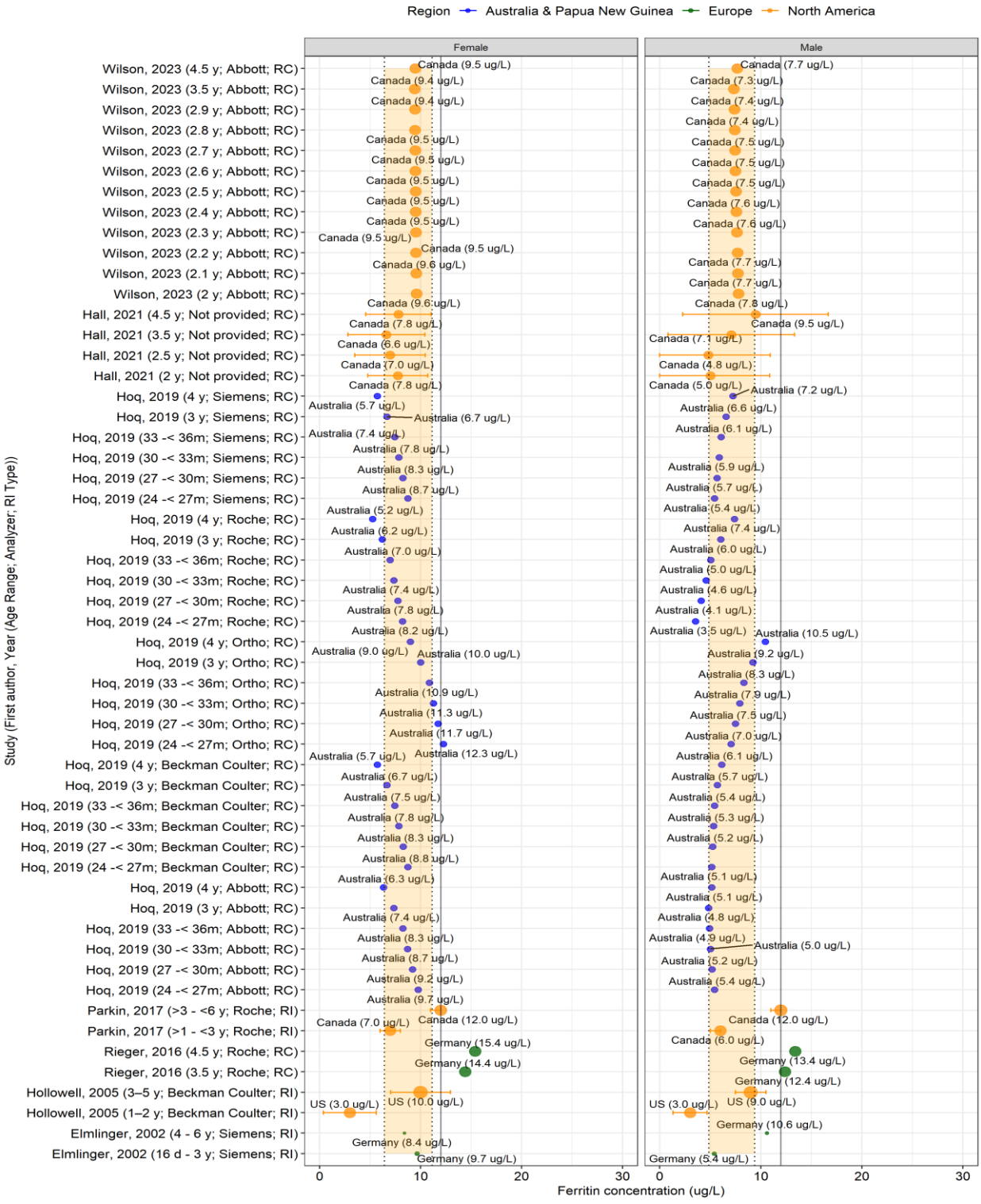


Figure S20. Distribution of ferritin RIs lower limits for 2-5 years old males and females

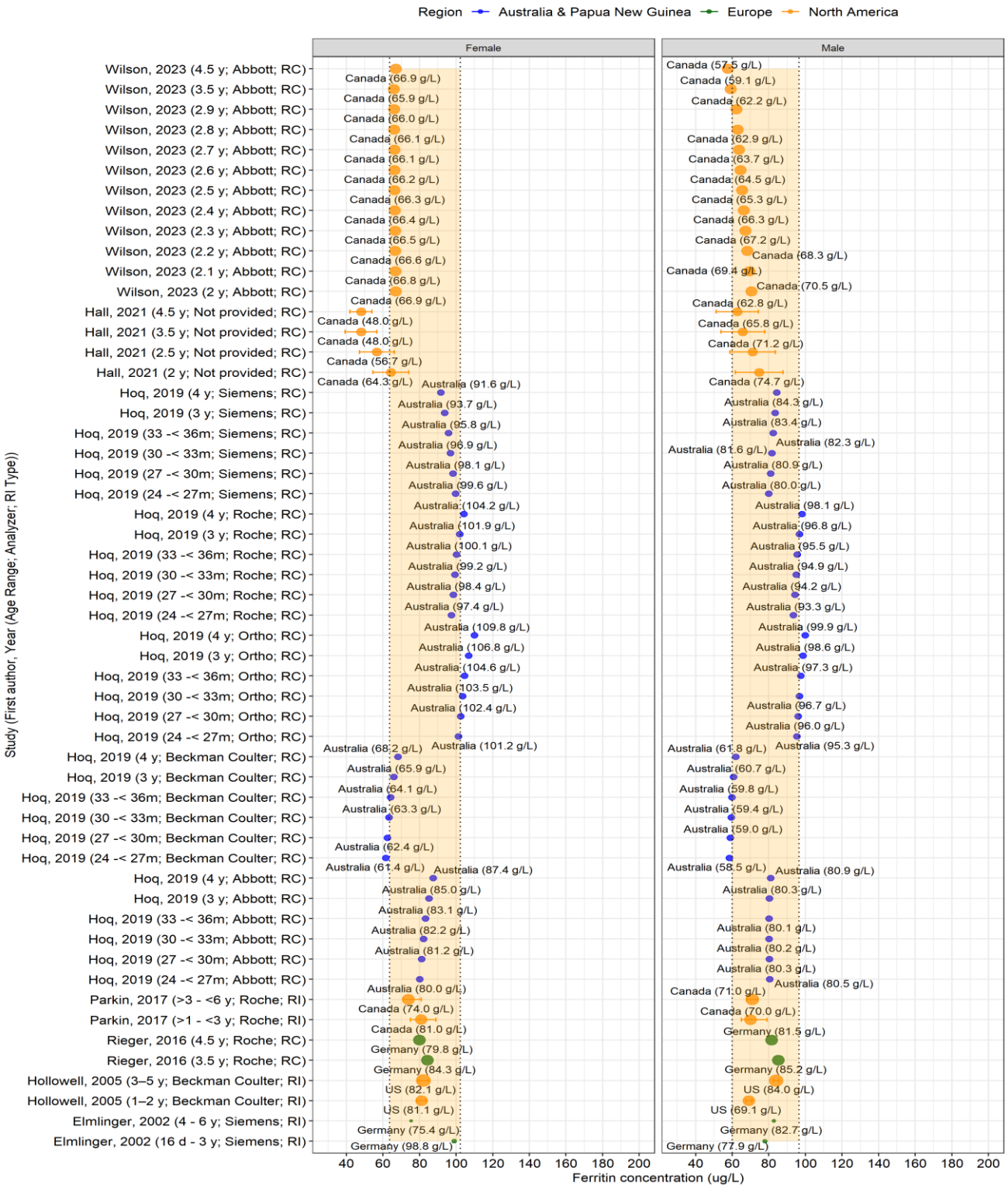


Figure S21. Distribution of ferritin RIs upper limits for 2-5 years old males and females

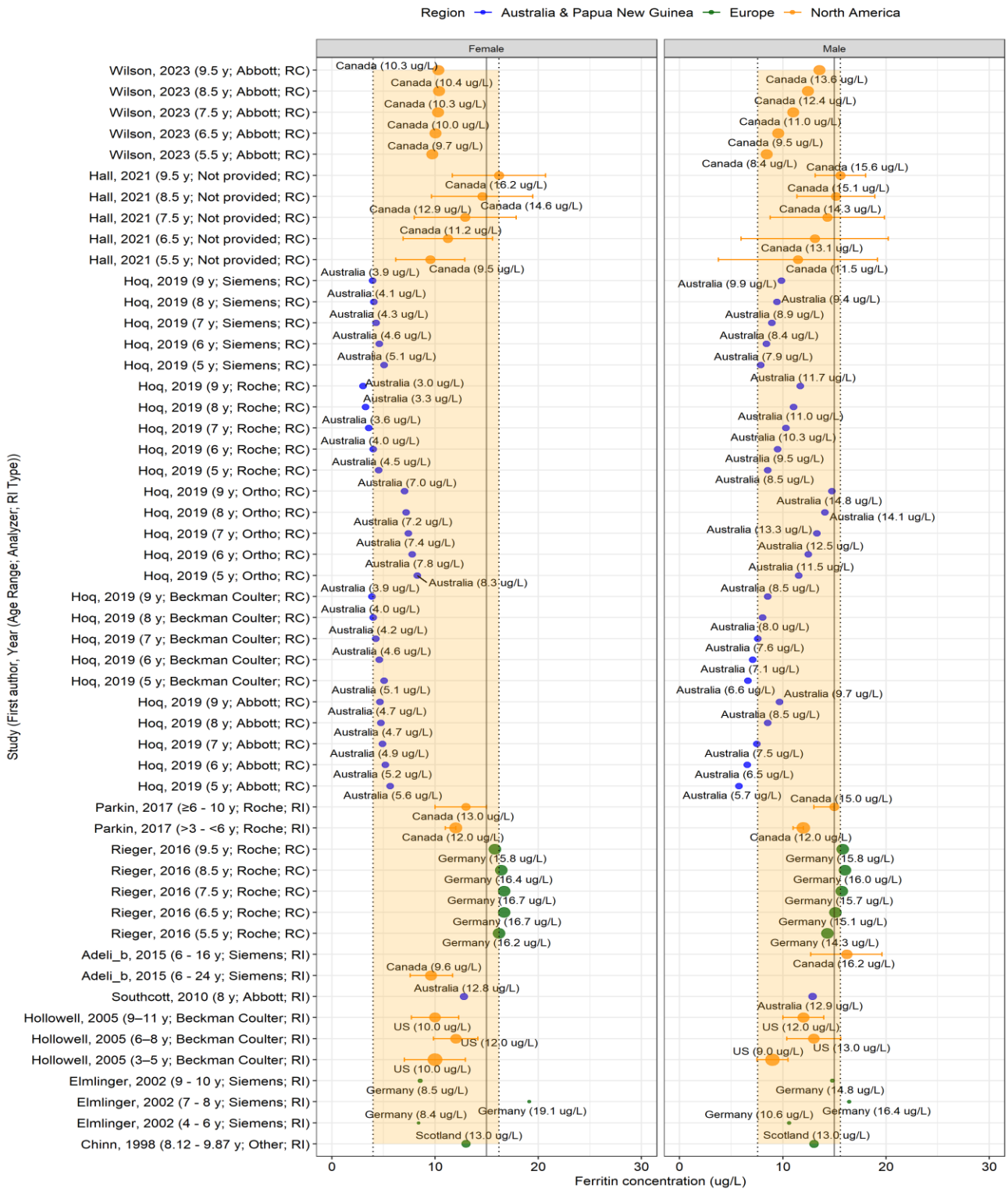


Figure S22. Distribution of ferritin RIs lower limits for 5-10 years old males and females

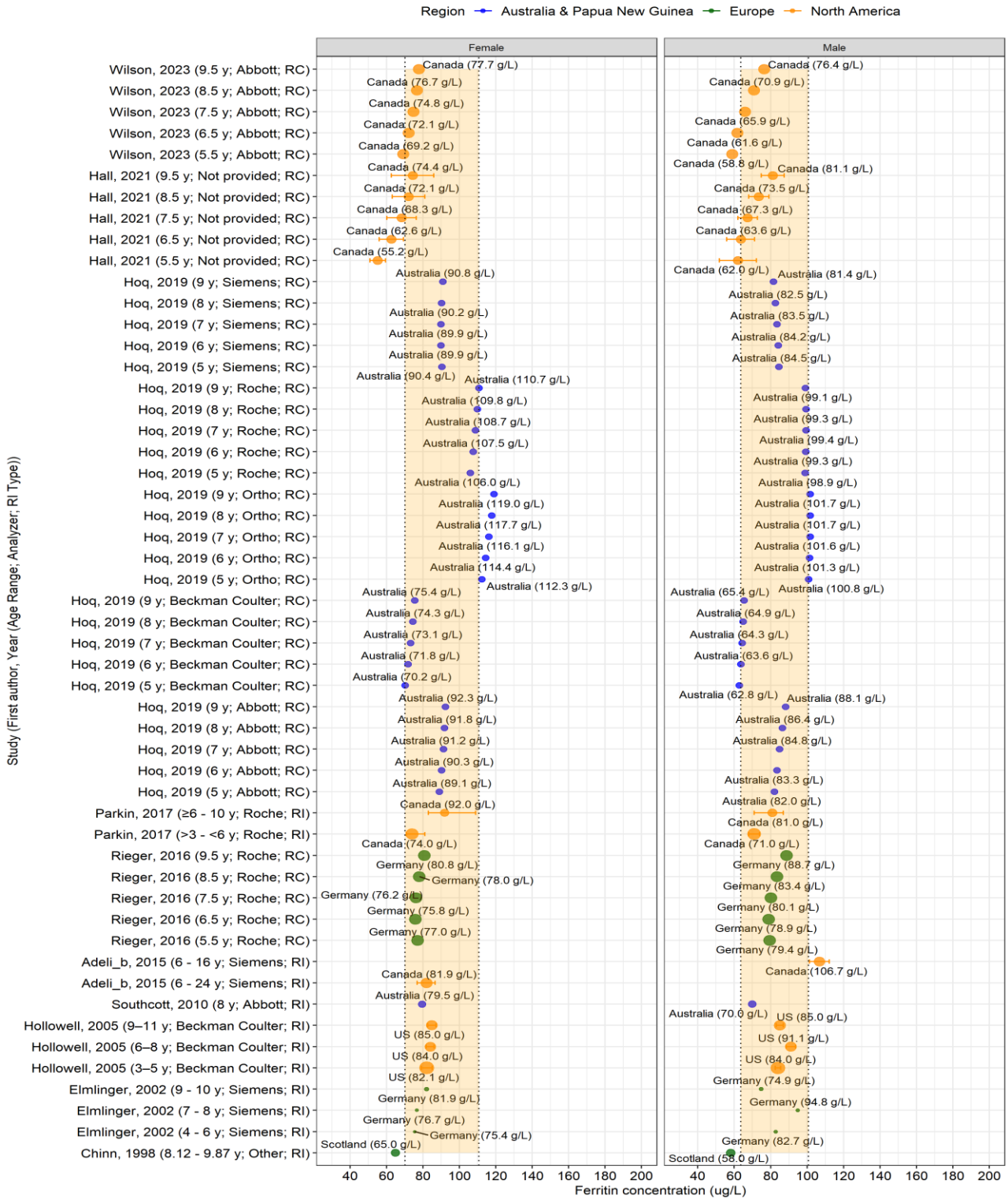


Figure S23. Distribution of ferritin RIs upper limits for 5-10 years old males and females

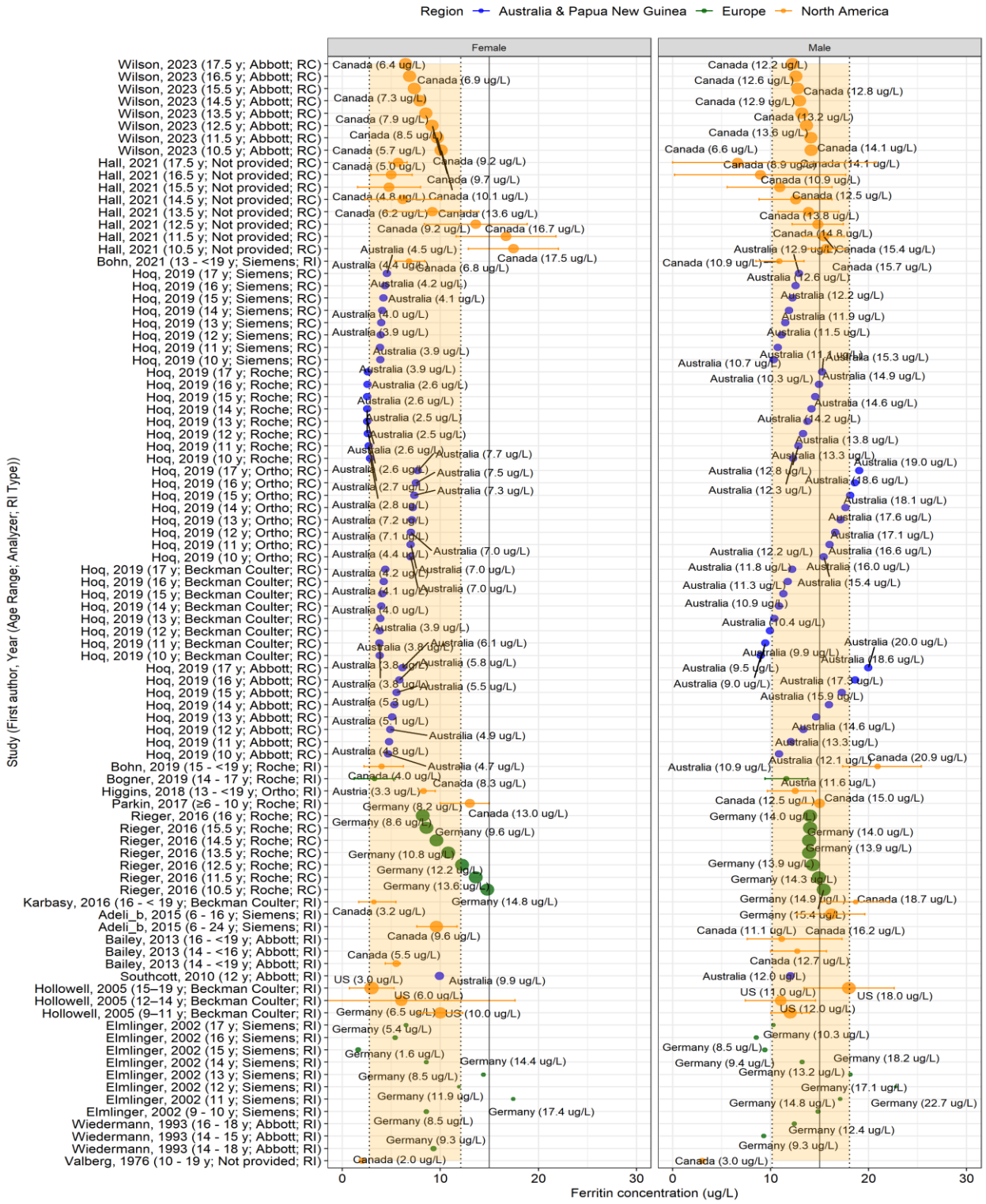


Figure S24. Distribution of ferritin RIs lower limits for 10-18 years old males and females

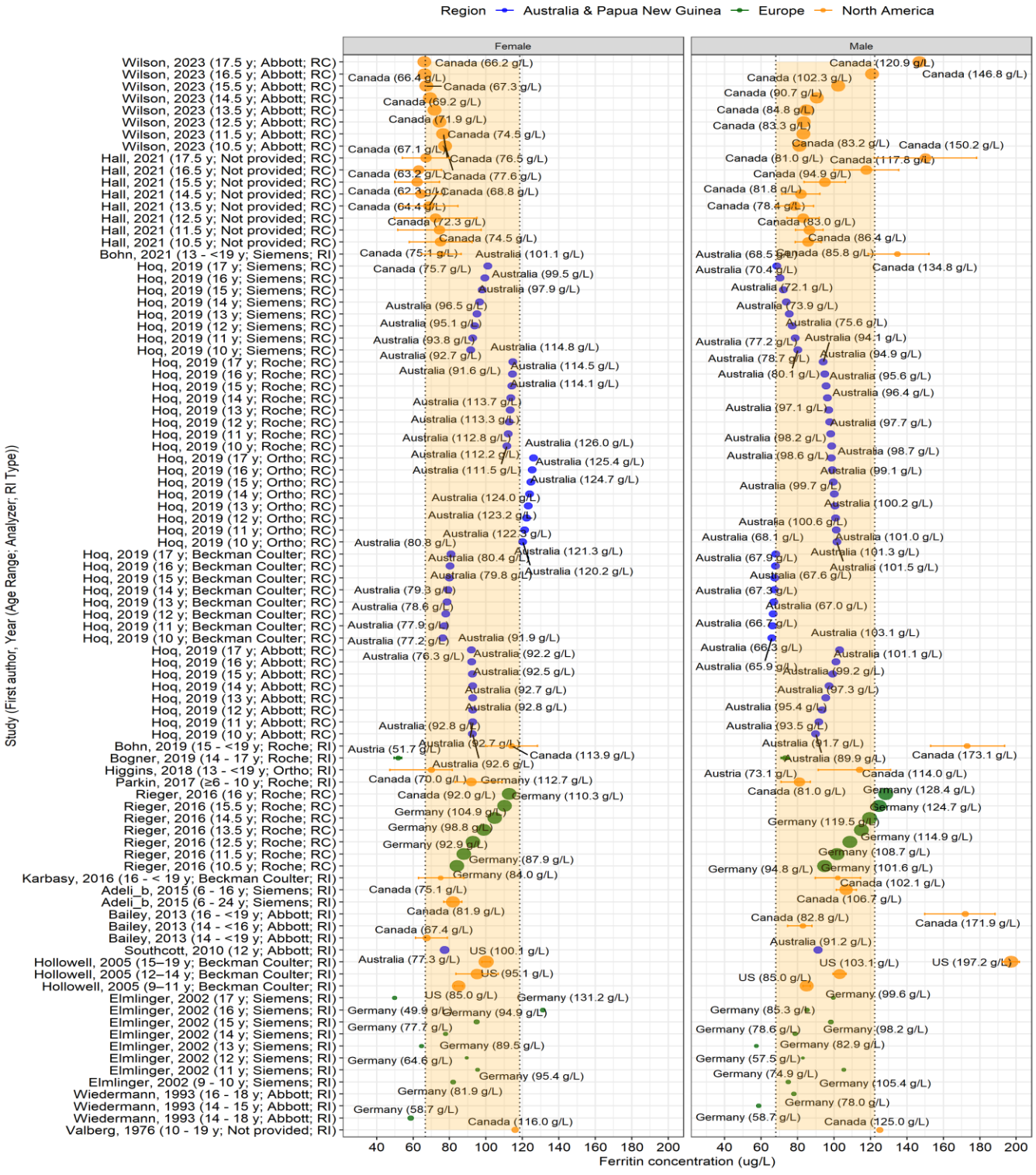


Figure S25. Distribution of ferritin RIs upper limits for 10-18 years old males and females

CHAPTER 5: BEYOND REFERENCE INTERVALS: OPTIMIZING LOWER AND UPPER LIMITS FOR REPORTING NORMATIVE HEMOGLOBIN LEVELS FOR CHILDREN

5.1. Preface to Chapter 5

Chapter 5 presents the fourth manuscript of this thesis, addressing the second objective of our research. In this manuscript, we establish reference curves for hemoglobin based on a cohort of healthy children in Canada. In addition to conventional reference intervals and curves, we develop *optimal* reference curves. This manuscript also outlines a methodological framework for estimating reference curves using data from cohort studies. In addition, we created an interactive Shiny web application that allows a user to obtain both traditional and optimal hemoglobin intervals for a child of any age between 2-weeks and 10.99 years old (Optimized Reference Assessment for Clinical Laboratory Evaluation of Hemoglobin [ORACLE-H]: https://svijetric.shinyapps.io/ORACLE-H/). The web tool also supports a multi-measure assessment approach, allowing individual laboratory results to be evaluated simultaneously against traditional reference intervals, optimal intervals, and global WHO thresholds. The manuscript submitted for review to *Blood*.

VB led the study design, review and selection of optimality criteria, performed all analysis including estimation of reference and optimal interval and curves and missing data imputation, development of Shiny app, presentation and interpretation of results and has drafted the initial manuscript. BKP, PCP, and JSH were involved in the design, data interpretation, writing and revision of the manuscript. FM and ML were involved in the design, data interpretation and revision of the manuscript. CMB, CSB, and JLM were involved in data acquisition and revision of the manuscript.

5.2 Manuscript status: under review for publication to *JAMA Network Open*

Beyond reference intervals: optimizing lower and upper limits for reporting normative hemoglobin levels for children

Vid Bijelić[1,2,4], Franco Momoli[1], Mira Liebman[2], Beth K. Potter[1], Cornelia M. Borkhoff[4], Catherine S. Birken[4], Jonathon L. Maguire[4], Patricia C. Parkin[4,] Jemila S. Hamid[2,3,*]

1 School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
2 Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada
3 Department of Mathematics and Statistics, University of Ottawa, Ottawa, Ontario, Canada
4 The Hospital for Sick Children, Toronto, Ontario, Canada

* Corresponding Author:
Jemila S. Hamid
STEM Complex, 150 Louis-Pasteur Pvt
University of Ottawa
Ottawa, ON, K1N 6N5, Canada
jhamid@uottawa.ca

Abstract

Introduction: Clinicians traditionally rely on reference intervals for interpretation of their patients' laboratory test results. In the pediatric population, estimating reference intervals from population normative data requires partitioning with respect to age, sex and other important factors related to children's growth and development. This leads to limited sample sizes, which impacts the precision of the estimates. This study sought to address these limitations by using curve estimation, where age is incorporated as a continuous variable. To compliment the recently published World Health Organization (WHO) thresholds for hemoglobin, we also aimed to estimate optimal hemoglobin curves for children, which has the potential to inform more appropriate hemoglobin reporting standards.

Methods: We used cross-sectional data from an existing cohort of healthy Canadian children (TARGet Kids!) aged 11 years or younger to estimate reference and optimal hemoglobin curves. Reference curves were estimated using the complete population, while optimal curves excluded children with iron deficiency, prematurity, or low birthweight, those whose households were considered low-income based on defined thresholds, and those meeting criteria to be considered underweight or overweight. Nonparametric quantile regression with restricted cubic splines was used to create sex-specific reference and optimal hemoglobin curves, from which age-specific reference and optimal intervals can be derived. Missing data on these characteristics for defining optimal hemoglobin curves were estimated using multiple imputation by chained equations. Sensitivity analyses were conducted to assess the impact of missing data and different imputation strategies. We developed a web-based graphical and computational algorithm for visualization of the estimated reference and optimal hemoglobin curves and for calculating the reference and optimal limits for a user-specified age interval.

Results: Hemoglobin data were available from 2451 males and 2146 females aged 2 weeks to less than 11 years. Optimal lower limit curves were higher than the reference curves for females across all ages and for males up to two years of age. No meaningful differences were observed for upper limits. WHO hemoglobin thresholds were consistently higher than the optimal curve lower limits (2.5^{th} percentile curve) across all ages but only higher than the 5^{th} percentile curve in children 5-10 years. Narrowly defined sex- and age-specific reference and optimal intervals were calculated from the curves. Sensitivity analyses revealed minimal differences between imputation strategies, although complete-case analysis curves were below hemoglobin reference and optimal curves, especially in females.

Conclusion: Our study provides hemoglobin reference curves for the pediatric population, where age is used as a continuous variable, hence eliminating excessive age partitioning and mitigating the limitation of sample size. Our study also presents novel pediatric optimal hemoglobin curves by applying health-based criteria, employing robust statistical methods and comprehensively addressing missing data. The findings emphasize the clinical value of optimal curves over traditional reference intervals. Future research should aim to validate these curves in diverse populations and extend the age range studied.

Introduction

Hemoglobin is a critical biomarker in clinical practice, serving as an essential indicator for diagnosing and managing various hematological and systemic conditions. To effectively interpret hemoglobin levels in pediatric populations, clinicians rely on established reference intervals.[1] The Clinical Laboratory Standards Institute (CLSI) has developed guidelines for establishing reference intervals; the guidelines recommendations include the selection of a healthy reference population and outline methods and sample size requirements.[2] The interval includes the lower 2.5th percentile and the upper 97.5th percentile, with corresponding CLSI recommended 90% confidence intervals, according to sex (female/male) and specific age groupings. Following CLSI guidelines assists clinical laboratories to meet accreditation standards.[3]

Reference curves, which treat age as a continuous variable rather than defining finite partitions for estimation, have been shown to have distinct advantages over reference intervals, including highlighting the distributional characteristics of hemoglobin levels and facilitating more precise clinical decision-making.[1] This enhanced granularity is especially beneficial in pediatric settings where growth and development can influence hemoglobin levels.[1] We recently conducted a systematic review of pediatric reference intervals and curves for hemoglobin and identified 48 studies, of which only 13 reported reference intervals adherent to CLSI guidelines and only 2 reported reference curves.[4] We identified substantial heterogeneity related to age groupings, further supporting the advantage of continuous reference curves.[4]

In 2024, the World Health Organization (WHO) updated their guideline on defining anemia.[5] The guideline was based on an analysis by Braat et al of a pooled international sample of an optimally healthy population using specific clinical and laboratory exclusion criteria.[6] Sex-

specific thresholds for three pre-adolescent age groups (6-23 months, 24-59 months, 5-11 years) were calculated based on the 5th percentile of the hemoglobin distribution for each age group. A strength of the WHO hemoglobin thresholds for anemia is their basis in a rigorous definition of an optimally healthy population. However, the WHO thresholds provide only a lower limit based on the 5th percentile, whereas CLSI guidelines recommend that reference intervals provide a lower limit based on the 2.5th percentile and an upper limit based on the 97.5th percentile. Furthermore, the WHO thresholds apply to age groupings rather than treating age as a continuous variable.

Our aim was to incorporate the strengths of the approach taken by Braat et al to establish the WHO thresholds (i.e. include an optimally healthy population), with the standards developed by the CLSI to create intervals defined by 2.5th and 97.5th percentiles, using a curve-based approach (which we refer to as optimal curves) to address concerns related to small sample sizes. Our objectives were to: (i) create sex-specific reference curves and optimal curves for hemoglobin among pre-adolescent children; (ii) examine differences between reference and optimal curves in this context; (iii) examine the lower limit reference and optimal curves in relation to WHO hemoglobin thresholds; (iv) calculate reference intervals and optimal intervals using narrow age groupings; and (v) create a web-based interactive tool to facilitate the translation of our results into practice.

Methods

Study design and population

We used a cross-sectional design, relying on data collected from healthy children living in the Greater Toronto Area (GTA), Canada, who participated in the TARGet Kids! primary care study. TARGet Kids! is an ongoing longitudinal study started in 2008 where data is being collected during the participants' regular scheduled well-child/well-baby visits. In this study, pre COVID-

19 data collected from 2008 to 2020 was used.[7] These routine visits are accessible to almost all children in the province of Ontario who have a primary care provider, with costs covered by the publicly-funded provincial health insurance program.[7] Visits are scheduled to occur at 2 weeks, 2, 4, 6, 9, 12, 15, 18 months and then annually.[8] Parents who provided informed consent for their child to participate in TARGet Kids! were invited to complete an enrollment questionnaire that included information regarding the child's birthweight and maternal gestational age as well as other variables such as household income. Research assistants embedded in each participating primary care practice, who are trained in pediatric phlebotomy, collected anthropometric measures and a blood sample at multiple time points.[7] Collection of blood samples on multiple points was optional.[7]

For this study, data were available for children in the TARGet Kids! cohort at ages 2 weeks to 10 years. Exclusion criteria were: children with diagnosed health conditions affecting growth (for example, failure to thrive or cystic fibrosis), children with any additional diagnosed acute or chronic health conditions (other than asthma and high-functioning autism), children with severe developmental delay, and families unable to communicate in English.[7]

Sample collection and analysis

Blood samples were collected in lavender EDTA tubes and transported to the laboratory at Mount Sinai Services (MSS) the same day. MSS is accredited by Accreditation Canada Diagnostics to ISO 15189 Plus™ standards (https://www.sinaihealth.ca/areas-of-care/pathology-and-laboratory-medicine/licence-and-accreditation) and complies with regulatory standards (http://www.mountsinaiservices.com/). At the laboratory, blood samples were analyzed within 4–6 hours from collection. The Sysmex XN-9000 Hematology Analyzer (Japan) was used to

measure hemoglobin.[7] Serum ferritin and CRP were measured using a modular platform Roche Diagnostics (Switzerland).[7]

Optimality criteria

To create optimal curves and intervals, we pre-specified optimality criteria based on previous research[9–12] and recommendations for good iron status.[13–15] Based on these criteria, we excluded data from children with the following characteristics ascertained at the time of the blood collection for the hemoglobin value used in the optimality curve estimation: iron deficiency (serum ferritin <12 ug/L), living in low-income households (< CAD $40,000)[16], and under or overweight (zBMI+/-1.96*SD). zBMI was derived from height/length and weight measurements, collected by trained TARGet Kids! research assistant at each visit, used to calculate BMI (kg/m2), which was then standardized by age and sex based on WHO growth standards.[7,10,17] Data from children younger than 2 years of age were also excluded if the child had been premature (< 37 weeks gestational age) or low birthweight (< 2,500 grams).[15] Finally, data were excluded if the C-reactive protein was above 5 mg/L (which may falsely elevate ferritin).[13]

Reference and Optimal Hemoglobin Curve Estimation

We used cross-sectional data for RC estimation; therefore, repeated measurements from the same child were removed using a method we named “lean pick”. In a “lean pick” method if a child had a blood sample collected at multiple visits, data from the visit with the least covered age group was used. This allowed us to efficiently use the available data and overcome sparseness in some age groups.

Sex-specific reference and optimal hemoglobin curves, defined by 2.5th and 97.5th percentiles for hemoglobin across all ages, were estimated using non-parametric quantile regression with restricted cubic splines. The numbers of knots were chosen based on the

Bayesian information criterion (BIC).[18] Outliers were identified with adjusted Tukey's method.[19] Hemoglobin values that were 3 times the inter-quartile range (IQR) below the 25th percentiles or three times the IQR above the 75th percentiles were considered outliers, and were excluded prior to the analysis. We calculated 90% confidence intervals (CI) for the fitted percentiles, which were determined using 200 bootstrap percentile replicates. We also estimated the 5th percentile, with corresponding 90% CI, for comparative purposes with the WHO hemoglobin thresholds to define anemia.[6] Lower and upper limits of the reference and optimal intervals were calculated for standardized age intervals[4]: monthly, three-month age intervals for children below age 3 years, and 1 year age intervals for children from age 3. We also created a web-based computational tool using the R Shiny app Optimized Reference Assessment for Clinical Laboratory Evaluation of Hemoglobin (ORACLE-H), allowing the user to specify age (or age interval) and sex to generate the desired reference and optimal intervals. We used the *quantreg* package version 5.98 in the R software version 4.4.1 to perform the quantile regression.[20,21]

Handling of missing data

Prior to applying "lean-pick" method for cross-sectional data selection, we used multiple imputations by chained equations to address missing data for the optimality criteria. The method is available through *mice* package in the R statistical software and *miceadds* package.[22] Imputation was performed only for visits that occurred and specific measurements were not taken. In order to account for repeated measures within individuals, imputation for both time-varying variables (e.g., ferritin, BMI) and time-invariant variables (e.g., sex, prematurity) were incorporated using respectively *2l.pmm* and *2lonly.pmm mice* R package for imputation. Restricted cubic splines were used within imputation models to accommodate for nonlinearity in age and ferritin. The rms package in the R statistical software was used to implement restricted cubic spline (rcs).[23] The imputation model included age in months (as rcs), hemoglobin, ferritin

(as rcs), total household income (defined as the middle of the reporting interval, for example, 20,000-39,000 was defined as 30,000), zBMI score, prematurity, and low birthweight.

Description of the imputation strategies for the main and sensitivity analyses are shown on Figure 1. Our main imputation strategy followed an multiple imputation, then deletion (MID) approach, where missing data were imputed using the full dataset under the missing at random (MAR) assumption.[24] The full dataset includes individuals with missing hemoglobin values. After imputation, individuals with imputed hemoglobin values were excluded, hence our primary analysis uses only observed values of hemoglobin but incorporates imputed values of the optimality criteria variables. To assess robustness of the results, we conducted two sensitivity analyses: (1) a deletion then multiple imputation (DMI) approach[24], in which individuals with missing hemoglobin values were excluded prior to imputation; (2) a complete-case approach, retaining only fully observed records. Because the household income variable was missing in 19% percent of children in the sample, we conducted sensitivity analysis to assess differences in the optimality curves when low-income household status was excluded as an optimality criterion.

Based on the overall proportion of records with missing data, eighty-one imputed datasets were generated in IMD approach and twenty-one in DMI approach. The results from multiple imputed data sets were combined into final estimates based on Rubin rules.[25]

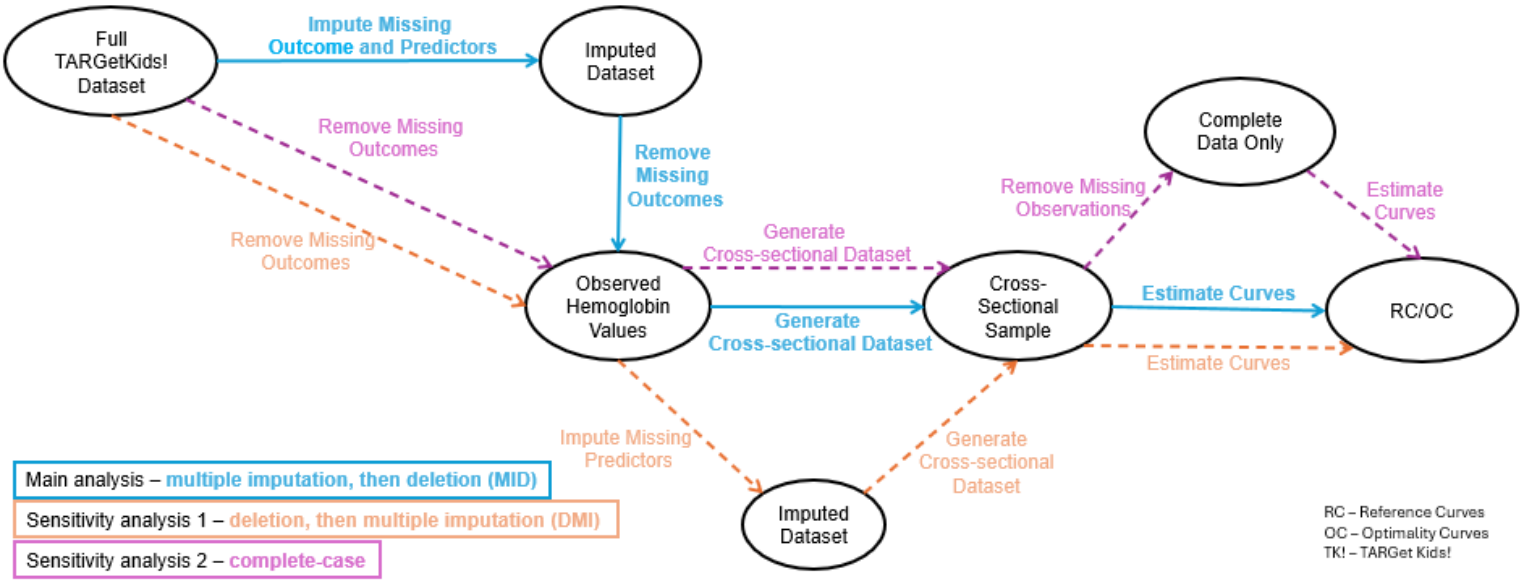


Figure 1. The primary analysis using multiple imputation, then deletion (MID) strategy to impute missing data and two sensitivity analysis using deletion, then multiple imputation (DMI) and complete-case analysis. Solid arrows represent the main imputation strategy path and dashed arrows represent sensitivity analysis strategies, colour coded as per legend. RC and OC represent reference and optimal hemoglobin curves, respectively.

Results

From 11857 children in TARGet Kids! enrolled from 2008 to 2020[7], samples from 2451 males and 2146 females, age 2 weeks to less than 11 years, were used to estimate sex-specific reference hemoglobin curves (Table 1). Samples from a subgroup of children meeting optimality criteria, including 1628 males and 1798 females, were used to estimate sex-specific optimal hemoglobin curves. Values from five males and four females were considered outliers and removed when estimating reference hemoglobin curves; values from two males were considered outliers and removed when estimating optimal hemoglobin curves.

Table 1. Characteristics of cross-sectional samples using "lean pick" method of hemoglobin in children 2 weeks to 10 years old*.

	MID sample for primary analysis* N = 4,597[1]	**Original TK! Sample* N = 11,802[1]**	**Complete case sample* N = 3,448[1]**	**DMI sample* N = 4,597[1]**	**Optimal Sample* N = 3,426[1]**
Characteristic					
Age in months, median (IQR)	38 (18, 63)	34 (11, 69)	36 (18, 61)	38 (18, 63)	45 (24, 68)
Sex, n (%)					
Female	2,146 (47)	5,672 (48)	1,611 (47)	2,146 (47)	1,628 (48)
Male	2,451 (53)	6,130 (52)	1,837 (53)	2,451 (53)	1,798 (52)
Hemoglobin (g/L), median (IQR)	121 (115, 127)	121 (116, 127)	121 (115, 126)	121 (115, 127)	121 (116, 127)
Missing		9,562			
Low income cut-off (LICO), n (%)					
Above LICO	4,151 (90)	7,688 (89)	3,110 (90)	4,148 (90)	3,426 (100)
Below LICO	446 (9.7)	941 (11)	338 (9.8)	449 (9.8)	
Missing		3,173			
Prematurity (applicable to the first 24 months), n (%)					
Not premature	4,405 (96)	9,850 (96)	3,280 (95)	4,395 (96)	3,426 (100)
Premature	192 (4.2)	417 (4.1)	168 (4.9)	202 (4.4)	
Missing		1,535			
Low birthweight (applicable to the first 24 months), n (%)					
Low birthweight	176 (3.8)	409 (3.8)	151 (4.4)	181 (3.9)	
Not low birthweight	4,421 (96)	10,285 (96)	3,297 (96)	4,416 (96)	3,426 (100)
Missing		1,108			
Obese/Underweight (Above or Below ZBMI +/- 1.96*SD), n (%)					
Not Obese/Underweight	4,231 (92)	10,313 (91)	3,192 (93)	4,236 (92)	3,426 (100)
Obese/Underweight	366 (8.0)	1,071 (9.4)	256 (7.4)	361 (7.9)	
Missing		418			
Iron deficiency (ID)**, n (%)					
ID	319 (6.9)	154 (6.4)	244 (7.1)	323 (7.0)	
No ID	4,278 (93)	2,264 (94)	3,204 (93)	4,274 (93)	3,426 (100)
Missing		9,384			

** All datasets were converted to cross-sectional format using the "lean pick" method. "MID sample for primary analysis" and "DMI sample" were based on the first imputed dataset generated from multiple imputation MID and DMI approaches respectively. "Optimal sample" was derived after applying predefined exclusion criteria on the first MID-imputed dataset. The "Original TARGet Kids! sample" includes the full analytic sample. "Complete case" sample consists of data after records with missing values were removed.*

*** ID was defined using WHO ferritin threshold[26]*

Sex-specific reference and optimal hemoglobin curves with corresponding 90% CIs are presented in Figure 2 and corresponding sample size across standardized age groups was

presented in Supplementary Table S2. For pre-adolescent females and for the first two years of life for males, the lower limits from optimal hemoglobin curves were higher than those from reference hemoglobin curves. There appeared to be no substantial differences between the upper limits of reference and optimal hemoglobin curves, for both females and males.

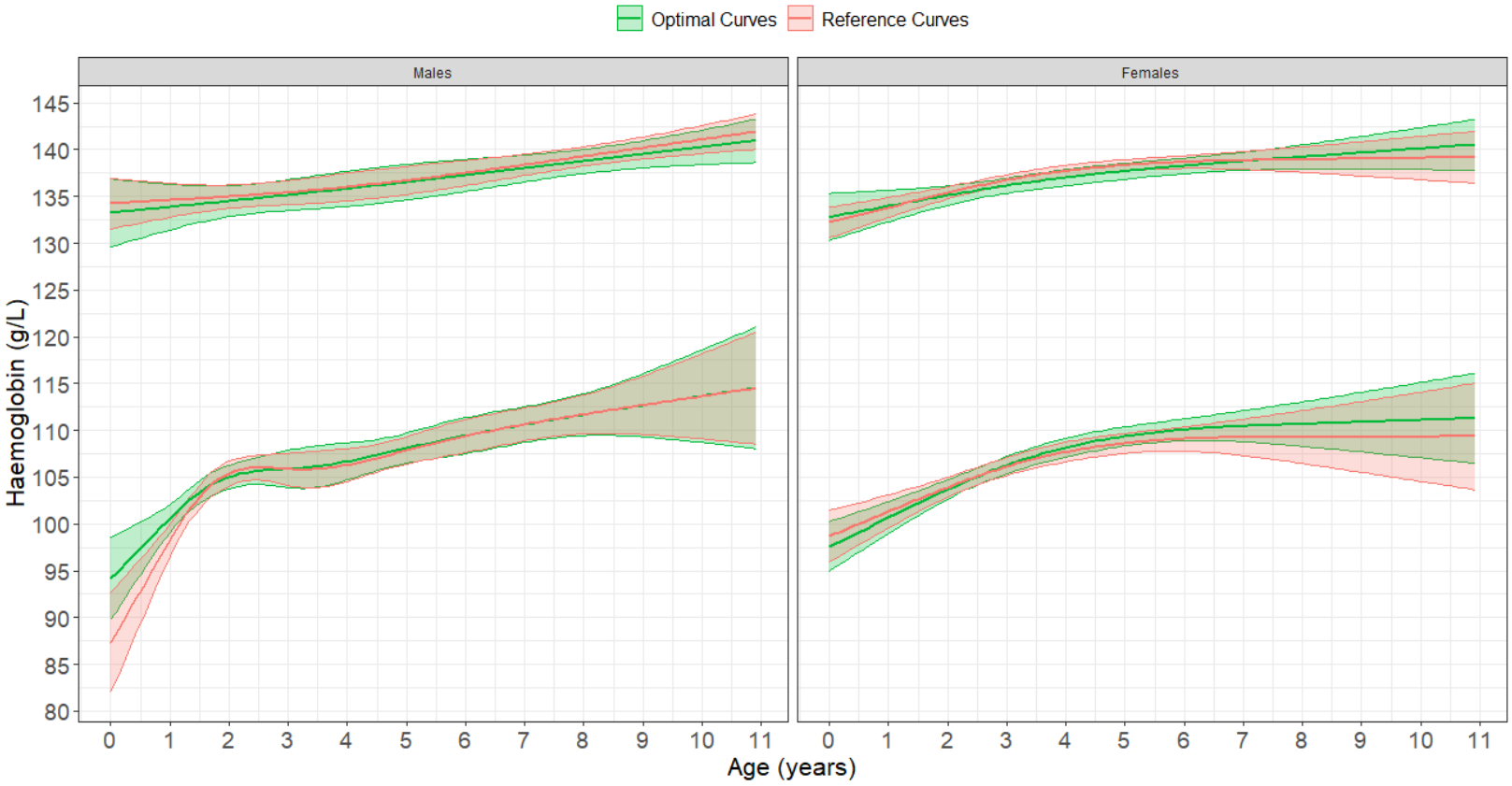


Figure 2. Male and female reference and optimal hemoglobin curves lower and upper limits with 90% confidence intervals.

Figure 3 shows the sex and age-specific WHO hemoglobin thresholds to define anemia [27] super-imposed on the optimal lower limit (2.5th percentiles) hemoglobin curves from our analysis, with the 5th, 10th, and 20th percentiles from our analysis also shown. The WHO hemoglobin thresholds were higher than the lower 2.5th percentile of the optimality hemoglobin

curves for all ages, but only higher than the 5th percentile for ages 5 to 10 years for both males and females.

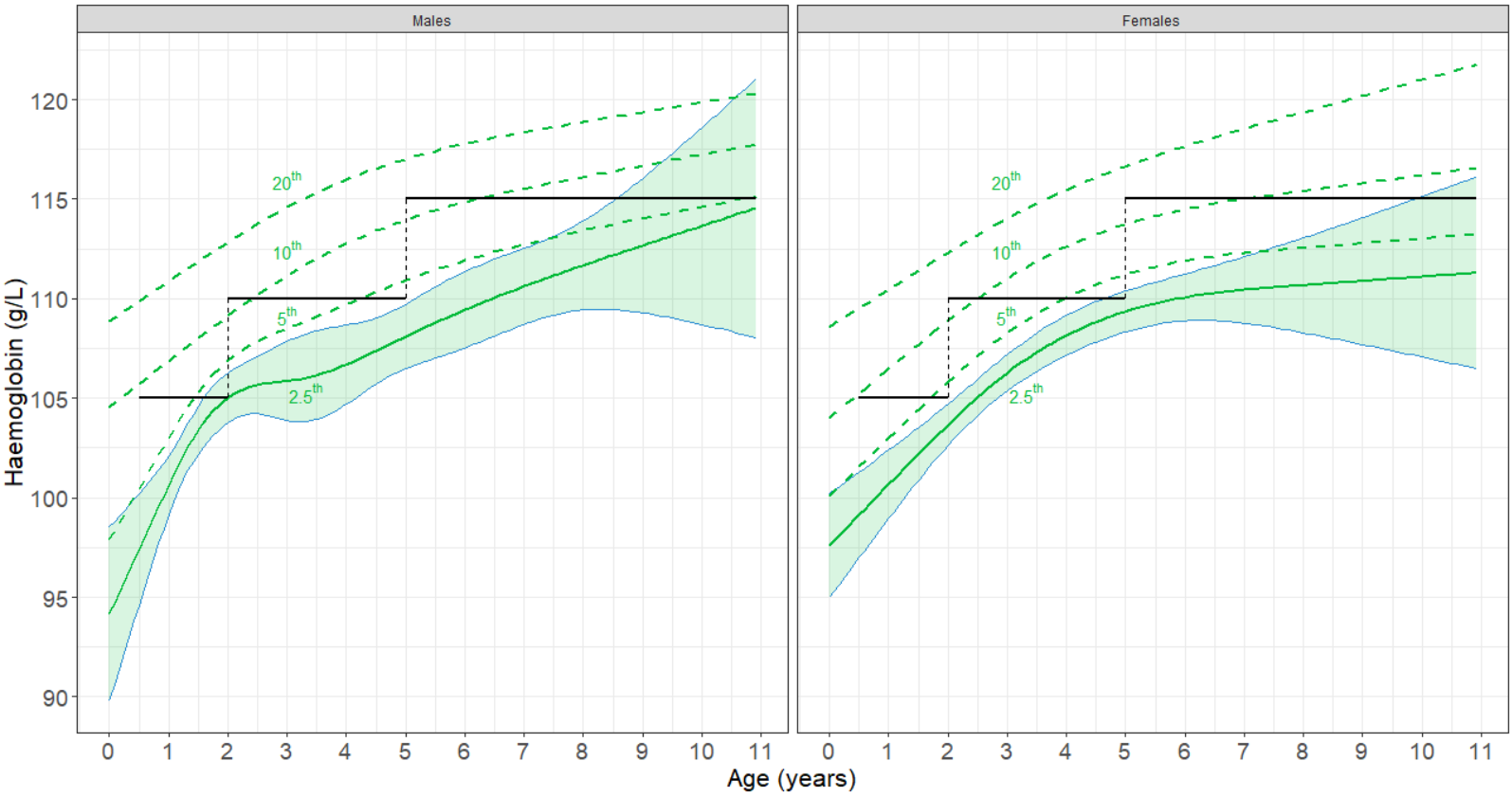


Figure 3. Male and female optimal hemoglobin curve with lower optimal limits (2.5th percentile) and WHO hemoglobin thresholds (solid black lines). Optimality 5th, 10th, and 20th, percentile hemoglobin curves were shown for orientation.

In Table 2, sex-specific reference and optimal intervals with the corresponding 90% CIs are provided for relatively narrow age partitions. For each month of age, from 0.5 to 131 months, sex-specific reference and optimal intervals with corresponding 90% confidence intervals are presented in Supplemental Table S1. Differences between reference and optimal intervals were similar to the differences described earlier for the differences between reference and optimal hemoglobin curves.

Table 2. Hemoglobin (g/L) sex-specific reference and optimal intervals with 90% confidence intervals averaged over narrow age groupings.

Age Group	Lower Limit				Upper Limit			
	Males Optimal	Males Reference	Females Optimal	Females Reference	Males Optimal	Males Reference	Females Optimal	Females Reference
2 weeks - 2 months	94.7 (90.6,98.8)	88.2 (83.2,93.3)	97.9 (95.3,100.4)	98.9 (96.3,101.6)	133.3 (129.8,136.8)	134.3 (131.6,136.9)	132.9 (130.5,135.4)	132.3 (130.7,133.9)
3 - 5 months	96.3 (93.0,99.6)	91.0 (86.9,95.1)	98.6 (96.3,100.9)	99.6 (97.2,102.0)	133.4 (130.2,136.6)	134.3 (131.9,136.8)	133.2 (131,135.4)	132.7 (131.3,134.2)
6 - 7 months	97.9 (95.4, 100.5)	93.7 (90.6,96.9)	99.4 (97.3,101.5)	100.2 (98.1,102.4)	133.6 (130.7,136.5)	134.4 (132.3,136.6)	133.5 (131.5,135.5)	133.1 (131.8,134.4)
9 - 11 months	99.6 (97.7,101.4)	96.5 (94.2,98.8)	100.2 (98.3,102.0)	100.9 (99.0,102.8)	133.8 (131.1,136.4)	134.5 (132.6,136.5)	133.8 (131.9,135.6)	133.5 (132.4,134.7)
12 - 14 months	101.1 (99.8,102.5)	99.2 (97.6,100.7)	100.9 (99.3,102.6)	101.6 (99.9,103.3)	133.9 (131.6,136.3)	134.6 (132.9,136.4)	134.1 (132.4,135.7)	133.9 (132.9,135)
15 - 17 months	102.6 (101.4,103.8)	101.6 (100.3,102.8)	101.7 (100.3,103.1)	102.2 (100.7,103.7)	134.1 (132,136.2)	134.7 (133.1,136.3)	134.4 (132.9,135.8)	134.3 (133.4,135.3)
18 - 20 months	103.7 (102.5,105.0)	103.5 (102.2,104.8)	102.5 (101.2,103.7)	102.9 (101.6,104.2)	134.2 (132.3,136.1)	134.8 (133.4,136.2)	134.6 (133.4,135.9)	134.7 (133.9,135.5)
21 - 23 months	104.6 (103.3,105.9)	104.8 (103.5,106.1)	103.2 (102.1,104.3)	103.5 (102.4,104.7)	134.4 (132.6,136.1)	134.9 (133.6,136.2)	134.9 (133.8,136.0)	135.1 (134.4,135.8)
24 - 26 months	105.2 (103.9,106.4)	105.6 (104.3,106.9)	103.9 (102.9,104.9)	104.1 (103.1,105.2)	134.5 (132.9,136.2)	135.0 (133.8,136.2)	135.2 (134.2,136.2)	135.5 (134.9,136.1)
27 - 29 months	105.5 (104.2,106.8)	105.9 (104.6,107.2)	104.6 (103.7,105.5)	104.7 (103.7,105.7)	134.7 (133.1,136.3)	135.1 (133.9,136.3)	135.5 (134.6,136.4)	135.9 (135.3,136.5)
30 - 32 months	105.7 (104.2,107.2)	106.0 (104.6,107.4)	105.3 (104.4,106.2)	105.3 (104.3,106.2)	134.9 (133.3,136.4)	135.2 (134,136.4)	135.7 (134.9,136.6)	136.2 (135.6,136.8)
33 - 35 months	105.8 (104.0,107.6)	105.9 (104.4,107.5)	105.9 (105.0,106.8)	105.8 (104.9,106.8)	135.0 (133.4,136.6)	135.3 (134.1,136.6)	136.0 (135.2,136.8)	136.5 (136,137.1)
3 - <4 years	106.2 (104.0,108.3)	105.9 (104.0,107.8)	107.2 (106.2,108.2)	106.9 (105.9,107.9)	135.4 (133.7,137.2)	135.7 (134.3,137.1)	136.6 (135.7,137.4)	137.2 (136.6,137.8)
4 - <5 years	107.3 (105.6,109.0)	107.0 (105.4,108.5)	108.8 (107.7,109.8)	108.2 (107.1,109.2)	136.1 (134.2,138.0)	136.3 (134.8,137.8)	137.3 (136.5,138.2)	138.0 (137.4,138.6)
5 - <6 years	108.7 (107.0,110.5)	108.6 (107.0,110.2)	109.7 (108.7,110.8)	108.9 (107.7,110.0)	136.8 (135.0,138.7)	137.1 (135.6,138.5)	138.0 (137.1,138.8)	138.5 (137.9,139.1)
6 - <7 years	110.0 (108.1,111.9)	110.0 (108.2,111.8)	110.3 (108.9,111.6)	109.2 (107.6,110.7)	137.6 (136.0,139.2)	137.9 (136.7,139.1)	138.5 (137.6,139.4)	138.8 (138.0,139.5)
7 - <8 years	111.1 (109.1,113.1)	111.1 (109.3,113.0)	110.6 (108.6,112.5)	109.3 (106.9,111.6)	138.3 (137.0,139.7)	138.8 (137.7,139.9)	139.0 (137.9,140.1)	138.9 (137.8,140.0)
8 - <9 years	112.1 (109.4,114.8)	112.2 (109.7,114.7)	110.8 (108.0,113.5)	109.3 (106.0,112.6)	139.1 (137.7,140.4)	139.7 (138.6,140.8)	139.4 (137.9,140.9)	139.0 (137.4,140.5)
9 - <10 years	113.1 (109.0,117.2)	113.1 (109.4,116.9)	111.0 (107.4,114.5)	109.3 (105.1,113.5)	139.8 (138.2,141.5)	140.6 (139.3,141.9)	139.9 (137.9,141.8)	139.1 (137,141.1)
10 - <11 years	114.1 (108.4,119.8)	114.1 (108.8,119.3)	111.2 (106.8,115.6)	109.3 (104.1,114.6)	140.6 (138.5,142.7)	141.5 (139.8,143.2)	140.3 (137.8,142.8)	139.1 (136.6,141.7)

We conducted sensitivity analysis to evaluate the impact of different imputation strategies on estimation of optimal hemoglobin curves. Figures 4 and 5 show optimal hemoglobin curves comparing three different imputation strategies: MID, DMI and complete case approach, separately for males and females. The MID and DMI methods (Supplemental Figure S1) show almost identical estimates for optimal hemoglobin curves across both sexes and for both lower and upper limits, suggesting minimal effect of imputation-deletion order on hemoglobin curve estimation.

Supplemental Figure S2 presents the comparison between MID approach (the primary analysis) and complete-case optimal hemoglobin curve estimation. The results show that complete-case estimation tends to underestimate hemoglobin levels for pre-school children and pre-adolescent females. The results also show greater variability for estimates from complete-data analysis compared to those based on imputed data. This result is consistent with the general characteristic of complete-case analysis, demonstrating that deleting observations with missing data, in most cases, is a less efficient approach than any of the established methods for handling missing data. Much less difference is observed for males.

As part of our sensitivity analysis, we removed family income as an optimality criterion, because of its large percentage of missingness. The results from this sensitivity analysis show no differences in the optimal hemoglobin curves (Supplemental Figures S3 and S4).

The ORACLE-H Shiny app (https://svijetric.shinyapps.io/ORACLE-H/) provides a user-friendly platform for visualizing the study's findings and exploring hemoglobin reference and optimal intervals in children. This tool is designed to specifically enhance clinical interpretation and evaluation of hemoglobin levels in pediatric patients; however, the approach can be used for any population and any biomarker. ORACLE-H allows clinicians to view sex- and age-specific

lower and upper reference/optimal limits for an individual child, based on their exact age. Users can also calculate reference and optimal intervals for custom-defined age ranges (age-partitions) within the age range of the study dataset (0 to 10 years). We also specifically incorporated WHO's age categories as well as the standardized age groups we introduced in a previous study to allow evidence synthesis[4] (these are 3-month intervals for children under 3 years and 1-year intervals for children older than 3 years). The app supports individualized assessment at a specific age (allowing the entry of a child's birth date or exact age in years, months, and days) by providing corresponding age- and sex-specific optimal and reference intervals. In addition, clinicians can input a hemoglobin measurement for a specific child, which will be visualized against optimality regions, reference hemoglobin curves, and WHO hemoglobin thresholds. Supplementary Figure S5 provides interface for ORACLE-H app.

Discussion

In this study, we aimed to move beyond traditional approaches to developing pediatric reference intervals for hemoglobin. We estimated sex-specific reference hemoglobin curves for the lower (2.5%) and upper (97.5%) limits from a large sample of healthy children 2 weeks to 10 years, recruited from primary care settings following CLSI guidelines. We then derived age interval specific reference intervals from these hemoglobin curves. We also applied pre-specified optimality criteria and created optimal hemoglobin curves. For the lower limits, optimal hemoglobin curves were higher than reference hemoglobin curves for early school age and pre-adolescent females; and for the first two years of life for males. Furthermore, when examined in relation to the 2024 WHO hemoglobin thresholds in three pre-adolescent age groups, WHO hemoglobin thresholds were often higher not only than the lower 2.5th but also 5th percentile for the optimal hemoglobin distribution. To provide a comprehensive visualization of the optimal

hemoglobin curves, we created a web-based tool, where the user can specify an individual child's age and sex and the application returns reference and optimal hemoglobin curves with customized age-specific intervals. Interpretation of hemoglobin levels may be improved when accompanied by lower and upper limits created using an age- and sex-specific curve-based approach and optimality criteria.

TARGet Kids! is an ongoing longitudinal observational cohort study and a certain amount of missing data is expected. In previous studies based on TARGet Kids! data, all missing data were assumed to be missing at random and multiple imputation was used for missing data.[10,12,28–31] A study by Maguire and colleagues (2013)[32] analyzed TARGet Kids! data using both complete case analysis, removing records with missing observations, as well as analysis based on imputed data. Although 24% of participants were missing survey, anthropometric, and lab measurements information, they found no difference between results using the two missing data strategies. In this study we analyzed data using MID approach, which included all available data in a missing data imputation model and then deleted records with imputed hemoglobin values prior to hemoglobin curve estimation. This approach was first introduced by von Hippel (2007)[24] for handling missing data. This approach protects from potentially poorly imputed outcome data, while still using maximum information from observed data. We also conducted sensitivity analysis to investigate how sensitive hemoglobin curve estimation is to difference imputation strategies. While the order of imputation had minimal effect on hemoglobin curve estimation, the complete-case analysis occasionally produced estimates that were lower than imputed model and wider confidence intervals. This observation contradicts the findings from the previous study that found no difference in imputed and complete case analysis results when answering a different question from this same dataset. While their conclusion supports

robustness of a complete case approach, our results highlights that such an approach may misrepresent important differences in the lower and upper extremes of data distribution.

We recently conducted a systematic review and meta-analysis of studies reporting pediatric reference intervals and curves for hemoglobin.[4] We identified 13 studies involving more than 35,000 children reporting on reference intervals adhering to CLSI guidelines.[4] Of these, six were from North America (Canada, USA), five from Asia (China, South Korea), and two from Africa (Ethiopia, Zimbabwe). Ten of the 13 studies used the Sysmex analyzer. Some studies reported on a restricted age group, for example one study focused on neonates[33], five studies on preschool and school aged children[8,34–37], and two studies on adolescents.[38,39] Most studies reported reference intervals for broad age groups. Only one study excluded children with laboratory evidence of iron deficiency (as defined by a low serum ferritin)[37]; however, this study only included children attending their 1-year medical check-up. For our meta-analysis, pooling of reference interval estimates was limited due to a high level of heterogeneity.[4] Sources of heterogeneity included age, sex and width of the age group.

The largest study used data from more than 20,000 children in the U.S. National Health and Nutrition Examination Survey (NHANES) to create reference intervals for children 1-19 years of age excluding children who reported medical conditions.[40] Using piecewise regression analysis, the authors reported an age-related increase in reference intervals for both males and females, with a decrease in the lower limit for females 15-19 years. The analysis yielded eight age categories for males (1-2 years, 3-5 years, 6 years, 7-11 years, 12 years, 13 years, 14-15 years, 16-19 years) and five age categories for females (1-3 years, 4 years, 5-6 years, 9-14 years, 15-19 years). However, this is the only study that used the Beckman Coulter analyzer and the authors note that transferability of their intervals need to be verified by local laboratories.

Furthermore, there was no data reported on children under 1 year of age. Except for the first two years of life, hemoglobin levels (male and female) we have estimated where lower than Fulgoni and colleagues.

In our systematic review, we identified two studies reporting on pediatric reference curves for hemoglobin by Wilson et al (Canada, n=496) and Yan et al (China, n=12,823).[41,42] In both studies, healthy children were recruited from hospital outpatient clinics and community sites, young infants were excluded (under 1 year in Wilson et al and under 3 months in Yan et al), and Sysmex was the analyzer used. Neither study excluded data on children with a low serum ferritin. While authors of both studies highlighted the advantages of reference curves over reference intervals, both noted current challenges of implementing reference curves into laboratory information systems.

We have identified some similarities and differences between our reference hemoglobin curves and the hemoglobin curves by Wilson et al and Yan et al.[41,42] Wilson et. al., (2021) observed a constant trend for females below age of 5 years, while Yan et. al., (2022) and our study observed an increasing trend during the same age period. Overall estimates in our study were generally lower than what was found in Yan et. al., (2022). Our hemoglobin curve and the hemoglobin curve from Yan et al show a downward trend in the lower limit in pre-adolescent and adolescent females; however, this is not shown in the curve from Wilson et al. Similar to our study, Yan et al. included children below one year of age; however, while our study included children as young as two weeks, their study was limited to those six months and older.

Strengths of our study include our novel approach to create optimal hemoglobin curves by applying optimality criteria, specifically excluding children with iron deficiency, living in low-income households, under or overweight, elevated C-reactive protein, and (if under 2 years

of age) prematurity or low birthweight. Additional strengths include our analytic approach which employed nonparametric quantile regression, utilizing restricted cubic splines as a smoothing function of age, handling missing data using multiple imputation strategies and conducting sensitivity analysis. The quantile regression method is independent of the underlying data distribution, enhancing its robustness against outliers and making it particularly advantageous when no suitable transformation is available to achieve normality.[20,43] Furthermore, the use of restricted cubic splines improves the model's ability to handle estimation in the tails of a distribution where data is usually sparse[23], which is common for pediatric populations.[1] Finally, our study underscores the value of performing sensitivity analysis and cautions against solely relying on complete case analysis when estimating lower and upper limits of reference and optimal hemoglobin curves.

Our study has some limitations. Data was collected from children in Toronto, Canada which may limit the generalizability of our findings to other populations. Exclusion of non-English-speaking families may limit the generalization of our study findings to these populations. In addition, we did not have sufficient data for children under 2 weeks of age and over 11 years of age. Moreover, while CLSI guideline suggests minimum of 120 sample for estimation of reference intervals using non-parametric method, there are no CLSI guidelines for sample size requirement for estimating reference curves. Further simulation studies similar to extensive reference intervals simulation studies conducted by Daly et al., (2017)[1] are needed to establish sample size requirements for reference curve estimation. Furthermore, for optimality criteria, while we excluded children with any known chronic disease, we did not have information to allow exclusion of children with asymptomatic hematologic conditions which may have been accompanied by a lower hemoglobin. Another limitation that may influence our

optimal curve is measurement error related to variables used to define optimality criteria. For example, parent reported birthweight may be subject to recall bias. Research conducted in North America[44,45] suggests it is a reasonably reliable alternative when hospital records are unavailable, while studies conducted in some Asian[46] settings have found that parent-reported low birthweight can underestimate the actual weight. Future studies may consider validating or directly obtaining this information from medical charts; however, this may require additional consent and data access permissions. In addition, while a previous study using TARGet Kids! data assumed that missingness was due completely at random and performed complete-case analysis[32], our comparison between imputed and complete-case results showed that, in addition to some differences appearing to be due to random variation and reduced sample size, we also observed systematic differences across specific age and sex combinations, indicating that the data were not missing completely at random. To assess the potential impact, we conducted a sensitivity analysis excluding the optimality criterion with the highest proportion of missingness, that is, socioeconomic status defines as household income. The results showed no meaningful differences in the estimated curves with or without this exclusion criterion applied. Finally, the large number of missing hemoglobin measurement may have influenced the estimation of reference and optimal hemoglobin curves. However, sensitivity analyses comparing imputation strategies using both observed and imputed hemoglobin values and using only observed values showed negligible differences in the estimated curves.

Conclusion

Our study presents novel pediatric optimal hemoglobin curves using TARGet Kids! cohort of healthy ethnically diverse Canadian children (Toronto area) ages 2-weeks up to 11 years, employing robust statistical methods and comprehensively addressing missing data. We showed

that for this population globally established WHO thresholds were often higher than lower 2.5th and 5th percentiles for the OCs indicating the need for revisiting currently recommended hemoglobin recommendations for these ages. Further research is needed to establish optimal hemoglobin curves for other iron related biomarkers such as serum and plasma ferritin for the same age group and extend the age range studied.

References

1 Daly CH, Higgins V, Adeli K, Grey VL, Hamid JS. Reference interval estimation: Methodological comparison using extensive simulations and empirical data. *Clin Biochem* 2017; **50**: 1145–58.

2 Clinical and Laboratory Standards Institute (CLSI), editor. Defining, Establishing and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition. Wayne, Pa: Clinical and Laboratory Standards Institute, 2008.

3 Peter TF, Rotz PD, Blair DH, Khine A-A, Freeman RR, Murtagh MM. Impact of Laboratory Accreditation on Patient Care and the Health System. *Am J Clin Pathol* 2010; **134**: 550–5.

4 Bijelić V, Bijelić M, Larock J, *et al.* Paediatric Reference Intervals and Curves for Haemoglobin Estimated Using Direct Methods: A Systematic Review and Meta-Analysis. *Int J Lab Hematol* 2025; : ijlh.14489.

5 World Health Organization, editor. Guideline on haemoglobin cutoffs to define anaemia in individuals and populations. Geneva: World Health Organization, 2024.

6 Braat S, Fielding KL, Han J, *et al.* Haemoglobin thresholds to define anaemia from age 6 months to 65 years: estimates from international data sources. *Lancet Haematol* 2024; **11**: e253–64.

7 Carsley S, Borkhoff CM, Maguire JL, *et al.* Cohort Profile: The Applied Research Group for Kids (TARGet Kids!). *Int J Epidemiol* 2015; **44**: 776–88.

8 Parkin PC, Hamid J, Borkhoff CM, *et al.* Laboratory reference intervals in the assessment of iron status in young children. *BMJ Paediatr Open* 2017; **1**: e000074.

9 Oatley H, Borkhoff CM, Chen S, *et al.* Screening for Iron Deficiency in Early Childhood Using Serum Ferritin in the Primary Care Setting. *Pediatrics* 2018; **142**: e20182095.

10 Sypes EE, Parkin PC, Birken CS, *et al.* Higher Body Mass Index Is Associated with Iron Deficiency in Children 1 to 3 Years of Age. *J Pediatr* 2019; **207**: 198-204.e1.

11 Borkhoff SA, Parkin PC, Birken CS, Maguire JL, Macarthur C, Borkhoff CM. Examining the Double Burden of Underweight, Overweight/Obesity and Iron Deficiency among Young Children in a Canadian Primary Care Setting. *Nutrients* 2023; **15**: 3635.

12 Bayoumi I, Parkin PC, Birken CS, Maguire JL, Borkhoff CM, for the TARGet Kids! Collaboration. Association of Family Income and Risk of Food Insecurity With Iron Status in Young Children. *JAMA Netw Open* 2020; **3**: e208603.

13 Baker RD, Greer FR, The Committee on Nutrition. Diagnosis and Prevention of Iron Deficiency and Iron-Deficiency Anemia in Infants and Young Children (0–3 Years of Age). *Pediatrics* 2010; **126**: 1040–50.

14 Domellöf M, Braegger C, Campoy C, *et al.* Iron Requirements of Infants and Toddlers. *J Pediatr Gastroenterol Nutr* 2014; **58**: 119–29.

15 Unger SL, Fenton TR, Jetty R, Critch JN, O'connor DL. Iron requirements in the first 2 years of life. *Paediatr Child Health* 2019; **24**: 555–555.

16 Statistics Canada. Low income cut-offs (LICOs) before and after tax by community size and family size, in current dollars. DOI:10.25318/1110024101-ENG.

17 Weltgesundheitsorganisation. Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age; methods and development. Geneva: WHO Press, 2006.

18 Schwarz G. Estimating the Dimension of a Model. *Ann Stat* 1978; **6**. DOI:10.1214/aos/1176344136.

19 Hubert M, Vandervieren E. An adjusted boxplot for skewed distributions. *Comput Stat Data Anal* 2008; **52**: 5186–201.

20 Koenker R. quantreg: Quantile Regression. 2024. https://CRAN.R-project.org/package=quantreg.

21 R Core Team. R: A Language and Environment for Statistical Computing. 2024. https://www.R-project.org/.

22 Buuren SV, Groothuis-Oudshoorn K. **mice** : Multivariate Imputation by Chained Equations in *R*. *J Stat Softw* 2011; **45**. DOI:10.18637/jss.v045.i03.

23 Harrell Jr FE. rms: Regression Modeling Strategies. 2024. https://CRAN.R-project.org/package=rms.

24 Von Hippel PT. 4. Regression with Missing Ys: An Improved Strategy for Analyzing Multiply Imputed Data. *Sociol Methodol* 2007; **37**: 83–117.

25 Buuren S van. Flexible imputation of missing data, Second edition. Boca Raton: Chapman & Hall/CRC, 2021.

26 World Health Organization. WHO guideline on use of ferritin concentrations to assess iron status in individuals and populations. Geneva: World Health Organization, 2020 https://iris.who.int/handle/10665/331505 (accessed June 18, 2025).

27 Pasricha S-R, Rogers L, Branca F, Garcia-Casal M-N. Measuring haemoglobin concentration to define anaemia: WHO guidelines. *The Lancet* 2024; **403**: 1963–6.

28 Cox KA, Parkin PC, Anderson LN, *et al.* Association Between Meat and Meat-Alternative Consumption and Iron Stores in Early Childhood. *Acad Pediatr* 2016; **16**: 783–91.

29 Saunders NR, Parkin PC, Birken CS, Maguire JL, Borkhoff CM. Iron status of young children from immigrant families. *Arch Dis Child* 2016; **101**: 1130–6.

30 Bayoumi I, Parkin PC, Lebovic G, *et al.* Iron deficiency among low income Canadian toddlers: a cross-sectional feasibility study in a Community Health Centre and non-Community Health Centre sites. *BMC Fam Pract* 2018; **19**: 161.

31 Gunaseelan V, Parkin PC, Wahi G, *et al.* Maternal ethnicity and iron status in early childhood in Toronto, Canada: a cross-sectional study. *BMJ Paediatr Open* 2020; **4**: e000635.

32 Maguire JL, Lebovic G, Kandasamy S, *et al.* The Relationship Between Cow's Milk and Stores of Vitamin D and Iron in Early Childhood. *Pediatrics* 2013; **131**: e144–51.

33 Li C-L, Tan C, Gao L, Chang Y, Fan W-M. Establishment and validation of haematological reference intervals for newborns aged 5 to 28 days in Nanjing, China. *Hematology* 2023; **28**: 2208891.

34 Hamid JS, Atenafu EG, Borkhoff CM, *et al.* Reference intervals for hemoglobin and mean corpuscular volume in an ethnically diverse community sample of Canadian children 2 to 36 months. *BMC Pediatr* 2021; **21**: 241.

35 Sun Q, Yue Y, Tian Y, Wang Q. Establishment of reference intervals for pediatric complete capillary blood counts: A multicenter study in Beijing. *Int J Lab Hematol* 2020; **42**: 533–43.

36 Wang G-C, Li N, Niu C, *et al.* Establishment of complete blood count reference intervals for Chinese preschoolers. *J Clin Lab Anal* 2017; **31**: e22095.

37 Lee HR, Shin S, Yoon JH, Roh EY, Chang JY. Reference Intervals of Hematology and Clinical Chemistry Analytes for 1-Year-Old Korean Children. *Ann Lab Med* 2016; **36**: 481–8.

38 Haileslasie H, Tsegaye A, Teklehaymanot G, *et al.* Community based hematological reference intervals among apparently healthy adolescents aged 12-17 years in Mekelle city, Tigrai, northern Ethiopia: A cross sectional study. *PLOS ONE* 2020; **15**: e0234106.

39 Gomani P, Matubu A, Mujuru H, *et al.* Hematological and Biochemical Laboratory Reference Intervals for Zimbabwean Adolescents. *Clin Lab* 2015; **61**. DOI:10.7754/Clin.Lab.2014.140506.

40 Fulgoni VL, Agarwal S, Kellogg MD, Lieberman HR. Establishing Pediatric and Adult RBC Reference Intervals With NHANES Data Using Piecewise Regression. *Am J Clin Pathol* 2019; **151**: 128–42.

41 Yan R, Peng Y, Hu L, *et al.* Continuous reference intervals for 21 biochemical and hematological analytes in healthy Chinese children and adolescents: The PRINCE study. *Clin Biochem* 2022; **102**: 9–18.

42 Wilson S, Bohn MK, Hall A, Higgins V, Abdelhaleem M, Adeli K. Continuous reference curves for common hematology markers in the CALIPER cohort of healthy children and adolescents on the Sysmex XN-3000 system. *Int J Lab Hematol* 2021; **43**: 1394–402.

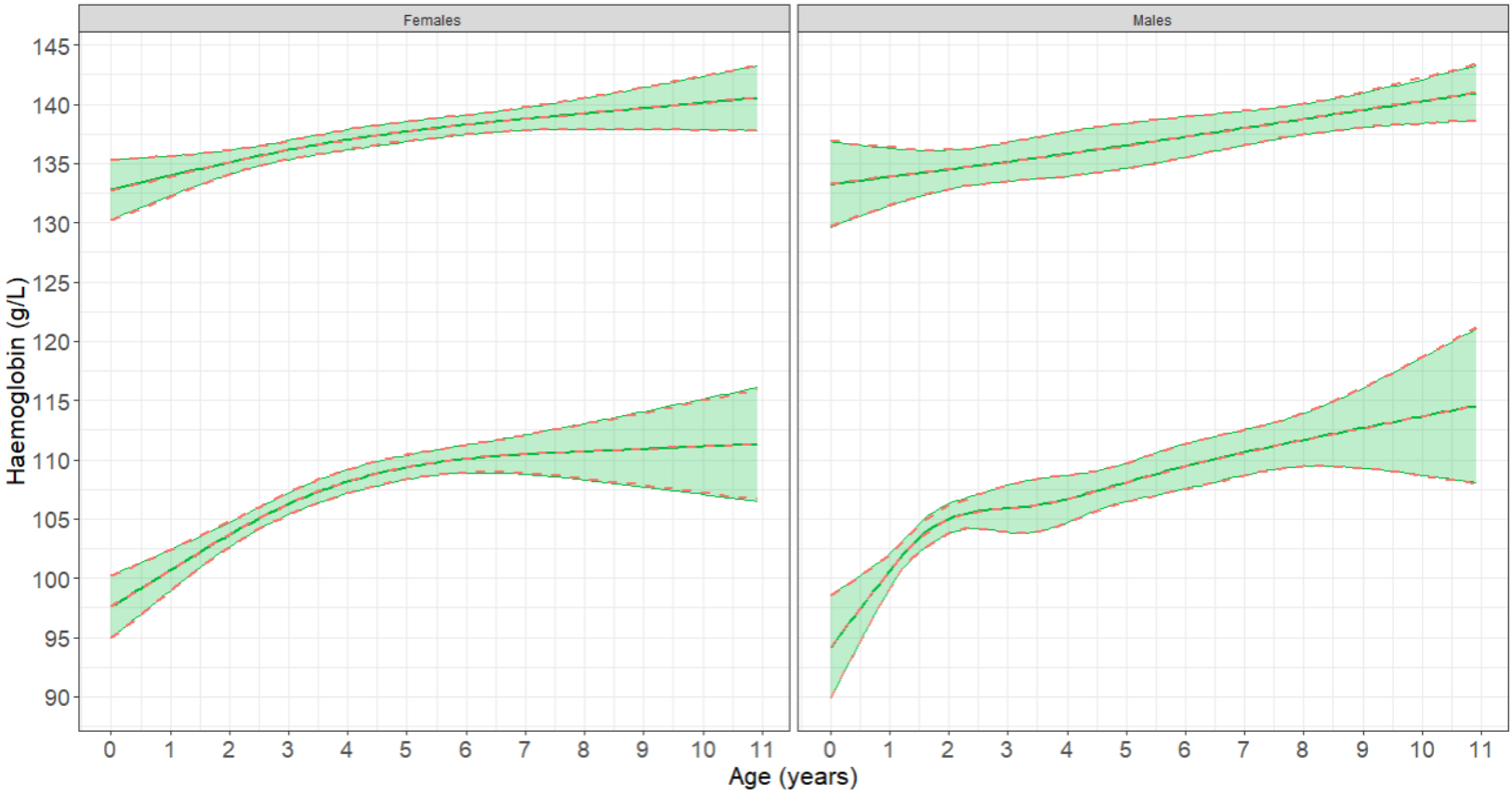
43 Koenker R, Bassett G. Regression Quantiles. *Econometrica* 1978; **46**: 33.

44 Lederman SA, Paxton A. Maternal Reporting of Prepregnancy Weight and Birth Outcome: Consistency and Completeness Compared with the Clinical Record. *Matern Child Health J* 1998; **2**: 123–6.

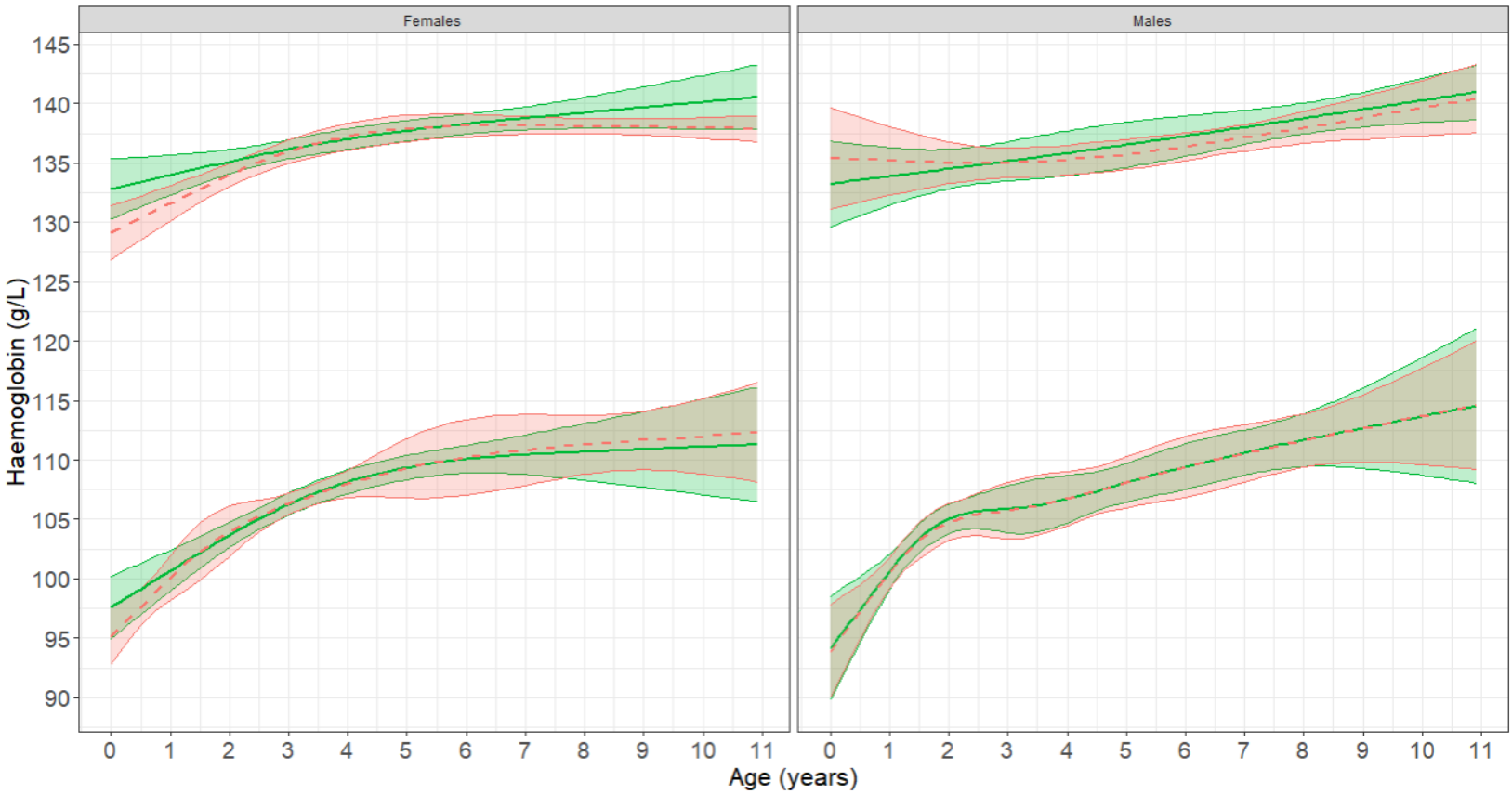
45 Hakim RB, Tielsch JM, See LC. Agreement between Maternal Interview- and Medical Record based Gestational Age. *Am J Epidemiol* 1992; **136**: 566–7.

46 Chang KT, Mullany LC, Khatry SK, LeClerq SC, Munos MK, Katz J. Validation of maternal reports for low birthweight and preterm birth indicators in rural Nepal. *J Glob Health* 2018; **8**: 010604.

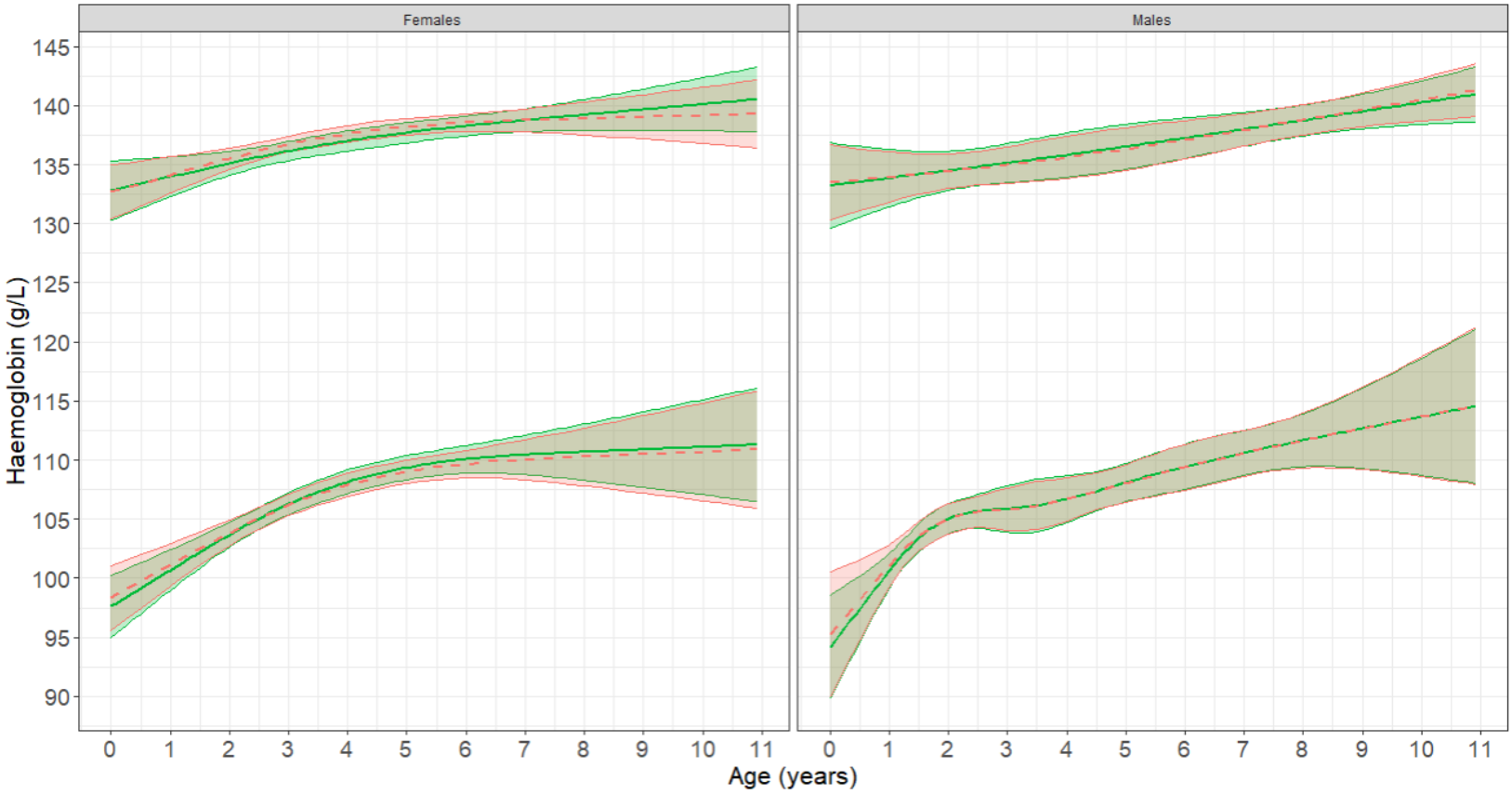
Supplemental tables and figures



Supplemental Figure S1. Optimal curves estimated using data from two different imputation strategies: multiple imputation, then deletion (MID, solid green lines) and deletion, then multiple imputation (DMI, dashed red). Curves are presented for both sexes. Ninety-percent confidence intervals for two strategies are shown in shaded green for MID and dashed red lines for DMI.

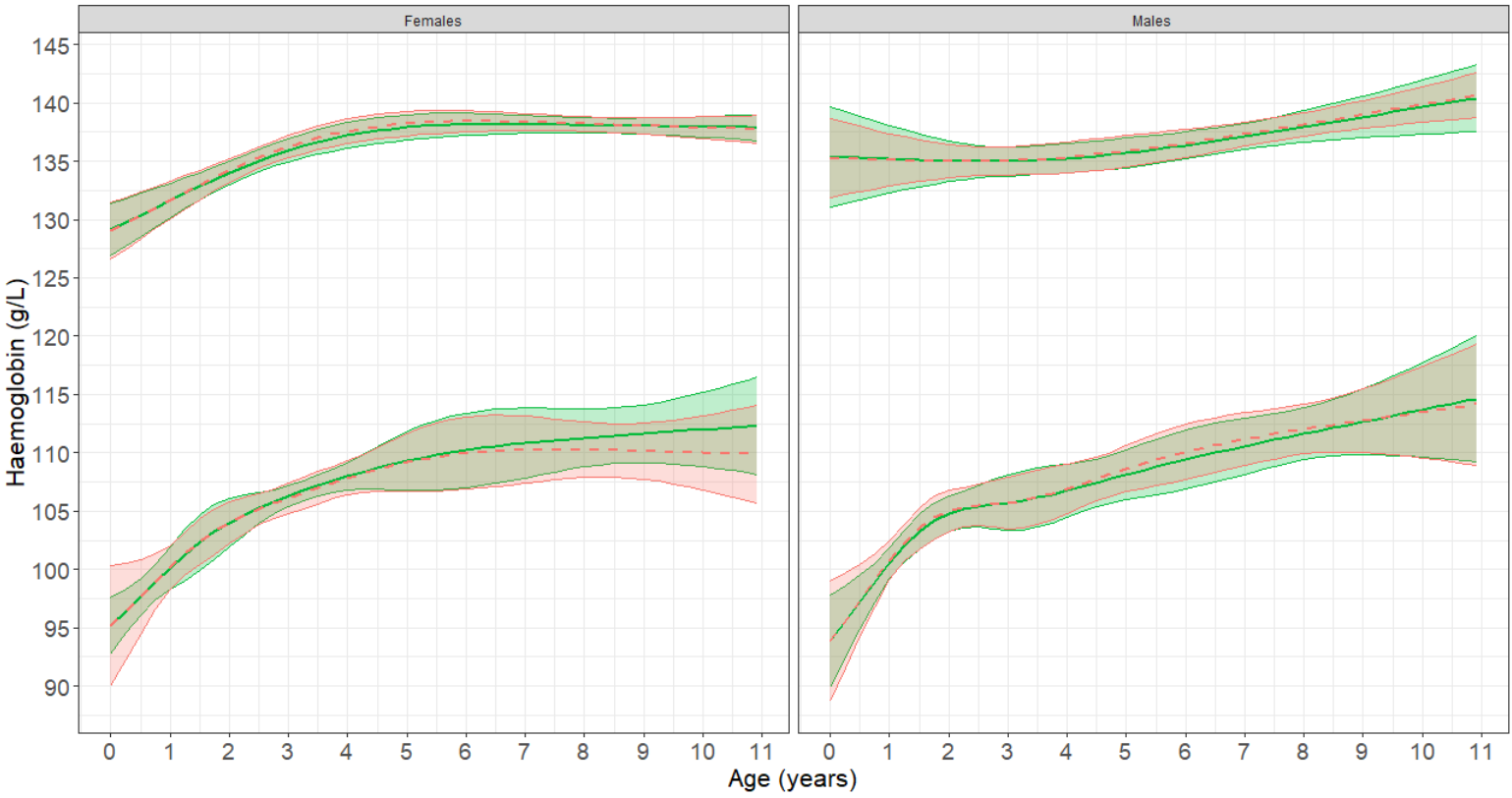


Supplemental Figure S2. Optimal curves estimated using data from two different imputation strategies: multiple imputation, then deletion (MID, solid green lines) and complete case (CC, dashed red). Curves are presented for both sexes. Ninety-percent confidence intervals for two strategies are shown in shaded green for MID and red for CC.

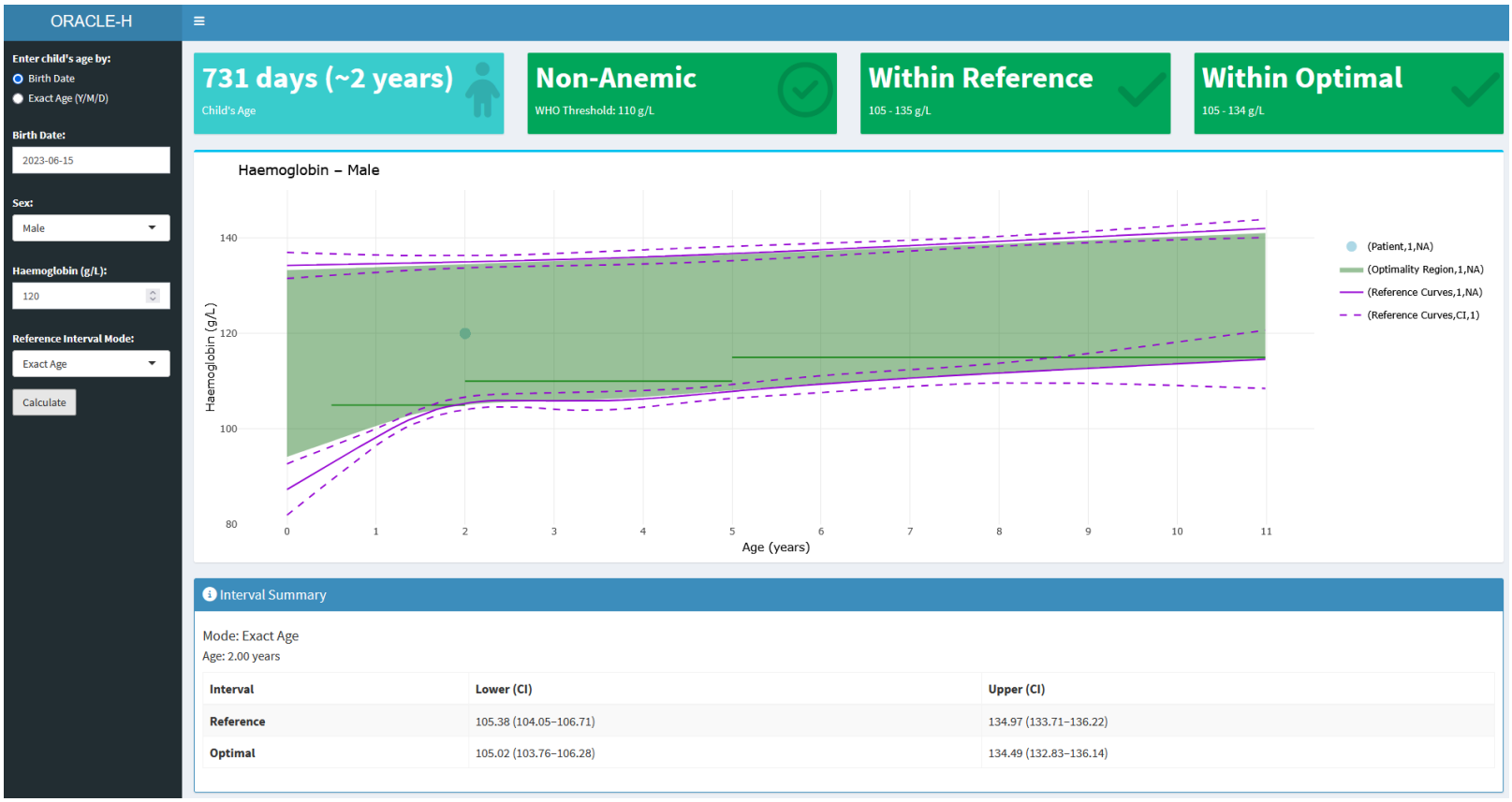


Supplemental Figure S3. Optimal curves estimated using data from multiple imputations, then deletion (MID, solid green lines) strategy and MID without LICO optimality criteria used

(dashed red line). Ninety-percent confidence intervals for two strategies are shown in transparent green for MID and red for MID without LICO.



Supplemental Figure S4. Optimal curves estimated using data from complete case analysis (solid green lines) strategy and complete case without LICO optimality criteria used (dashed red line). Ninety-percent confidence intervals for two strategies are shown in transparent green for complete case and red for complete case without LICO.



Supplemental Figure S5. Web-based Shiny app tool Optimized Reference Assessment for Clinical Laboratory Evaluation of Hemoglobin (ORACLE-H) for obtaining child's reference and optimal intervals at any age and WHO, standardized (3m/1y), and custom age intervals. Tools also allow for aiding hemoglobin level assessment based on multiple evaluation criteria. ORACLE-H can be accessed at: https://svijetric.shinyapps.io/ORACLE-H/

Supplemental Table S1. Sex-specific reference intervals (RI) and optimal intervals (OI) with 90% confidence intervals for 0 – 131 months of age.

	Lower Limits		Upper Limits	
Age (months)	**Reference Intervals**	**Optimal Intervals**	**Reference Intervals**	**Optimal Intervals**
Males				
0	87.3 (82.0-92.7)	94.2 (89.8-98.5)	134.2 (131.5-136.9)	133.2 (129.6-136.9)
1	88.2 (83.2-93.3)	94.7 (90.6-98.8)	134.3 (131.6-136.9)	133.3 (129.8-136.8)
2	89.2 (84.4-93.9)	95.2 (91.4-99.1)	134.3 (131.7-136.8)	133.3 (129.9-136.7)
3	90.1 (85.7-94.5)	95.8 (92.2-99.4)	134.3 (131.8-136.8)	133.4 (130.1-136.7)
4	91.0 (86.9-95.1)	96.3 (93.0-99.6)	134.3 (131.9-136.8)	133.4 (130.2-136.6)
5	91.9 (88.1-95.7)	96.9 (93.8-99.9)	134.4 (132.0-136.7)	133.5 (130.4-136.6)
6	92.8 (89.4-96.3)	97.4 (94.6-100.2)	134.4 (132.2-136.7)	133.5 (130.5-136.6)
7	93.7 (90.6-96.9)	97.9 (95.4-100.5)	134.4 (132.3-136.6)	133.6 (130.7-136.5)
8	94.7 (91.8-97.5)	98.5 (96.2-100.8)	134.5 (132.4-136.6)	133.6 (130.8-136.5)
9	95.6 (93.0-98.1)	99.0 (97.0-101.1)	134.5 (132.5-136.5)	133.7 (131.0-136.4)
10	96.5 (94.2-98.8)	99.6 (97.7-101.4)	134.5 (132.6-136.5)	133.8 (131.1-136.4)
11	97.4 (95.4-99.4)	100.1 (98.5-101.7)	134.6 (132.7-136.4)	133.8 (131.3-136.3)
12	98.3 (96.5-100.1)	100.6 (99.2-102.1)	134.6 (132.8-136.4)	133.9 (131.4-136.3)
13	99.2 (97.6-100.7)	101.1 (99.8-102.5)	134.6 (132.9-136.4)	133.9 (131.6-136.3)
14	100.0 (98.6-101.4)	101.6 (100.4-102.9)	134.6 (133.0-136.3)	134.0 (131.7-136.2)
15	100.8 (99.5-102.1)	102.1 (100.9-103.3)	134.7 (133.0-136.3)	134.0 (131.8-136.2)
16	101.6 (100.3-102.8)	102.6 (101.4-103.8)	134.7 (133.1-136.3)	134.1 (132.0-136.2)
17	102.3 (101.0-103.5)	103.0 (101.8-104.2)	134.7 (133.2-136.3)	134.1 (132.1-136.2)
18	102.9 (101.7-104.2)	103.4 (102.2-104.6)	134.8 (133.3-136.2)	134.2 (132.2-136.1)
19	103.5 (102.2-104.8)	103.8 (102.5-105.0)	134.8 (133.4-136.2)	134.2 (132.3-136.1)
20	104.0 (102.7-105.3)	104.1 (102.8-105.3)	134.8 (133.5-136.2)	134.3 (132.4-136.1)

21	104.4 (103.1-105.8)	104.4 (103.1-105.6)	134.9 (133.5-136.2)	134.3 (132.5-136.1)
22	104.8 (103.5-106.1)	104.6 (103.3-105.9)	134.9 (133.6-136.2)	134.4 (132.6-136.1)
23	105.1 (103.8-106.5)	104.8 (103.6-106.1)	134.9 (133.7-136.2)	134.4 (132.7-136.1)
24	105.4 (104.1-106.7)	105.0 (103.8-106.3)	135.0 (133.7-136.2)	134.5 (132.8-136.1)
25	105.6 (104.3-106.9)	105.2 (103.9-106.4)	135.0 (133.8-136.2)	134.5 (132.9-136.2)
26	105.7 (104.4-107.0)	105.3 (104.0-106.6)	135.0 (133.8-136.3)	134.6 (133.0-136.2)
27	105.9 (104.6-107.2)	105.4 (104.1-106.7)	135.1 (133.9-136.3)	134.6 (133.1-136.2)
28	105.9 (104.7-107.2)	105.5 (104.2-106.8)	135.1 (133.9-136.3)	134.7 (133.1-136.3)
29	106.0 (104.7-107.3)	105.6 (104.2-107.0)	135.1 (133.9-136.3)	134.8 (133.2-136.3)
30	106.0 (104.7-107.3)	105.7 (104.2-107.1)	135.2 (134.0-136.4)	134.8 (133.2-136.4)
31	106.0 (104.6-107.4)	105.7 (104.2-107.2)	135.2 (134.0-136.4)	134.9 (133.3-136.4)
32	106.0 (104.6-107.4)	105.7 (104.1-107.4)	135.3 (134.0-136.5)	134.9 (133.3-136.5)
33	106.0 (104.5-107.5)	105.8 (104.1-107.5)	135.3 (134.1-136.5)	135.0 (133.4-136.6)
34	105.9 (104.4-107.5)	105.8 (104.0-107.6)	135.3 (134.1-136.6)	135.0 (133.4-136.6)
35	105.9 (104.2-107.5)	105.8 (103.9-107.7)	135.4 (134.1-136.6)	135.1 (133.5-136.7)
36	105.9 (104.1-107.6)	105.9 (103.9-107.8)	135.4 (134.1-136.7)	135.1 (133.5-136.8)
37	105.8 (104.0-107.6)	105.9 (103.8-108.0)	135.5 (134.2-136.8)	135.2 (133.5-136.9)
38	105.8 (104.0-107.6)	105.9 (103.8-108.1)	135.5 (134.2-136.8)	135.2 (133.6-136.9)
39	105.8 (103.9-107.7)	106.0 (103.8-108.2)	135.5 (134.2-136.9)	135.3 (133.6-137.0)
40	105.8 (103.9-107.7)	106.0 (103.8-108.2)	135.6 (134.2-137.0)	135.4 (133.6-137.1)
41	105.8 (103.9-107.8)	106.1 (103.9-108.3)	135.6 (134.3-137.0)	135.4 (133.7-137.2)
42	105.9 (103.9-107.8)	106.2 (103.9-108.4)	135.7 (134.3-137.1)	135.5 (133.7-137.2)
43	105.9 (104.0-107.8)	106.2 (104.0-108.4)	135.7 (134.3-137.2)	135.5 (133.7-137.3)
44	106.0 (104.0-107.9)	106.3 (104.1-108.5)	135.8 (134.4-137.2)	135.6 (133.8-137.4)
45	106.0 (104.1-107.9)	106.4 (104.3-108.5)	135.8 (134.4-137.3)	135.6 (133.8-137.5)

46	106.1 (104.3-108.0)	106.5 (104.4-108.6)	135.9 (134.4-137.4)	135.7 (133.8-137.5)
47	106.2 (104.4-108.0)	106.6 (104.5-108.6)	135.9 (134.5-137.4)	135.8 (133.9-137.6)
48	106.3 (104.5-108.0)	106.7 (104.7-108.7)	136.0 (134.5-137.5)	135.8 (133.9-137.7)
49	106.4 (104.7-108.1)	106.8 (104.9-108.7)	136.1 (134.6-137.5)	135.9 (134.0-137.8)
50	106.5 (104.9-108.2)	106.9 (105.0-108.8)	136.1 (134.6-137.6)	135.9 (134.0-137.8)
51	106.6 (105.0-108.2)	107.0 (105.2-108.8)	136.2 (134.7-137.7)	136.0 (134.1-137.9)
52	106.8 (105.2-108.3)	107.1 (105.4-108.9)	136.2 (134.7-137.7)	136.0 (134.1-138.0)
53	106.9 (105.4-108.4)	107.2 (105.5-108.9)	136.3 (134.8-137.8)	136.1 (134.2-138.0)
54	107.0 (105.6-108.5)	107.4 (105.7-109.0)	136.3 (134.8-137.9)	136.2 (134.2-138.1)
55	107.2 (105.7-108.6)	107.5 (105.9-109.1)	136.4 (134.9-137.9)	136.2 (134.3-138.1)
56	107.3 (105.9-108.7)	107.6 (106.0-109.2)	136.5 (134.9-138.0)	136.3 (134.4-138.2)
57	107.4 (106.0-108.9)	107.7 (106.1-109.3)	136.5 (135.0-138.0)	136.3 (134.4-138.3)
58	107.6 (106.2-109.0)	107.8 (106.3-109.4)	136.6 (135.1-138.1)	136.4 (134.5-138.3)
59	107.7 (106.3-109.2)	108.0 (106.4-109.6)	136.6 (135.1-138.2)	136.5 (134.5-138.4)
60	107.9 (106.4-109.3)	108.1 (106.5-109.7)	136.7 (135.2-138.2)	136.5 (134.6-138.4)
61	108.0 (106.5-109.5)	108.2 (106.6-109.9)	136.8 (135.3-138.3)	136.6 (134.7-138.5)
62	108.1 (106.6-109.6)	108.3 (106.7-110.0)	136.8 (135.4-138.3)	136.6 (134.8-138.5)
63	108.3 (106.7-109.8)	108.4 (106.8-110.1)	136.9 (135.4-138.4)	136.7 (134.8-138.6)
64	108.4 (106.8-110.0)	108.6 (106.8-110.3)	137.0 (135.5-138.4)	136.8 (134.9-138.6)
65	108.5 (106.9-110.1)	108.7 (106.9-110.4)	137.0 (135.6-138.5)	136.8 (135.0-138.7)
66	108.7 (107.0-110.3)	108.8 (107.0-110.6)	137.1 (135.7-138.5)	136.9 (135.1-138.7)
67	108.8 (107.1-110.4)	108.9 (107.1-110.7)	137.2 (135.7-138.6)	136.9 (135.1-138.7)
68	108.9 (107.2-110.6)	109.0 (107.2-110.9)	137.2 (135.8-138.6)	137.0 (135.2-138.8)
69	109.0 (107.3-110.7)	109.1 (107.3-111.0)	137.3 (135.9-138.7)	137.1 (135.3-138.8)
70	109.1 (107.4-110.9)	109.2 (107.3-111.1)	137.4 (136.0-138.8)	137.1 (135.4-138.9)

71	109.3 (107.5-111.0)	109.3 (107.4-111.2)	137.4 (136.1-138.8)	137.2 (135.5-138.9)
72	109.4 (107.6-111.1)	109.4 (107.5-111.3)	137.5 (136.2-138.9)	137.2 (135.5-139.0)
73	109.5 (107.7-111.3)	109.5 (107.6-111.5)	137.6 (136.3-138.9)	137.3 (135.6-139.0)
74	109.6 (107.8-111.4)	109.6 (107.7-111.6)	137.7 (136.3-139.0)	137.4 (135.7-139.0)
75	109.7 (107.9-111.5)	109.7 (107.8-111.7)	137.7 (136.4-139.0)	137.4 (135.8-139.1)
76	109.8 (108.0-111.6)	109.8 (107.9-111.8)	137.8 (136.5-139.1)	137.5 (135.9-139.1)
77	109.9 (108.2-111.7)	109.9 (108.0-111.9)	137.9 (136.6-139.1)	137.6 (136.0-139.1)
78	110.0 (108.3-111.8)	110.0 (108.1-112.0)	137.9 (136.7-139.2)	137.6 (136.0-139.2)
79	110.1 (108.4-111.9)	110.1 (108.2-112.1)	138.0 (136.8-139.2)	137.7 (136.1-139.2)
80	110.2 (108.5-112.0)	110.2 (108.3-112.2)	138.1 (136.9-139.3)	137.7 (136.2-139.3)
81	110.3 (108.6-112.1)	110.3 (108.4-112.3)	138.2 (137.0-139.3)	137.8 (136.3-139.3)
82	110.4 (108.7-112.2)	110.4 (108.5-112.3)	138.2 (137.1-139.4)	137.9 (136.4-139.3)
83	110.5 (108.8-112.3)	110.5 (108.6-112.4)	138.3 (137.2-139.5)	137.9 (136.5-139.4)
84	110.6 (108.9-112.4)	110.6 (108.7-112.5)	138.4 (137.2-139.5)	138.0 (136.6-139.4)
85	110.7 (109.0-112.5)	110.7 (108.8-112.6)	138.4 (137.3-139.6)	138.1 (136.6-139.5)
86	110.8 (109.1-112.6)	110.8 (108.9-112.7)	138.5 (137.4-139.6)	138.1 (136.7-139.5)
87	110.9 (109.1-112.7)	110.9 (109.0-112.8)	138.6 (137.5-139.7)	138.2 (136.8-139.6)
88	111.0 (109.2-112.8)	111.0 (109.0-112.9)	138.7 (137.6-139.8)	138.2 (136.9-139.6)
89	111.1 (109.3-112.9)	111.1 (109.1-113.0)	138.7 (137.7-139.8)	138.3 (137.0-139.7)
90	111.2 (109.4-113.0)	111.2 (109.2-113.1)	138.8 (137.8-139.9)	138.4 (137.0-139.7)
91	111.3 (109.4-113.1)	111.2 (109.2-113.2)	138.9 (137.8-139.9)	138.4 (137.1-139.8)
92	111.4 (109.5-113.2)	111.3 (109.3-113.4)	139.0 (137.9-140.0)	138.5 (137.2-139.8)
93	111.5 (109.5-113.4)	111.4 (109.3-113.5)	139.0 (138.0-140.1)	138.6 (137.2-139.9)
94	111.5 (109.6-113.5)	111.5 (109.4-113.6)	139.1 (138.1-140.2)	138.6 (137.3-139.9)
95	111.6 (109.6-113.6)	111.6 (109.4-113.8)	139.2 (138.2-140.2)	138.7 (137.4-140.0)

96	111.7 (109.6-113.8)	111.7 (109.4-113.9)	139.3 (138.2-140.3)	138.7 (137.4-140.0)
97	111.8 (109.7-113.9)	111.8 (109.5-114.1)	139.3 (138.3-140.4)	138.8 (137.5-140.1)
98	111.9 (109.7-114.1)	111.8 (109.5-114.2)	139.4 (138.4-140.5)	138.9 (137.6-140.2)
99	112.0 (109.7-114.2)	111.9 (109.5-114.4)	139.5 (138.4-140.6)	138.9 (137.6-140.2)
100	112.0 (109.7-114.4)	112.0 (109.5-114.5)	139.6 (138.5-140.6)	139.0 (137.7-140.3)
101	112.1 (109.7-114.5)	112.1 (109.5-114.7)	139.6 (138.6-140.7)	139.1 (137.7-140.4)
102	112.2 (109.7-114.7)	112.2 (109.5-114.9)	139.7 (138.6-140.8)	139.1 (137.8-140.5)
103	112.3 (109.7-114.9)	112.3 (109.4-115.1)	139.8 (138.7-140.9)	139.2 (137.8-140.5)
104	112.4 (109.7-115.1)	112.3 (109.4-115.3)	139.9 (138.8-141.0)	139.2 (137.9-140.6)
105	112.4 (109.6-115.2)	112.4 (109.4-115.5)	139.9 (138.8-141.1)	139.3 (137.9-140.7)
106	112.5 (109.6-115.4)	112.5 (109.4-115.7)	140.0 (138.9-141.2)	139.4 (138.0-140.8)
107	112.6 (109.6-115.6)	112.6 (109.3-115.8)	140.1 (138.9-141.3)	139.4 (138.0-140.9)
108	112.7 (109.6-115.8)	112.7 (109.3-116.0)	140.2 (139.0-141.4)	139.5 (138.0-141.0)
109	112.8 (109.5-116.0)	112.8 (109.3-116.3)	140.2 (139.0-141.5)	139.6 (138.1-141.0)
110	112.8 (109.5-116.2)	112.8 (109.2-116.5)	140.3 (139.1-141.5)	139.6 (138.1-141.1)
111	112.9 (109.5-116.4)	112.9 (109.2-116.7)	140.4 (139.1-141.6)	139.7 (138.1-141.2)
112	113.0 (109.4-116.6)	113.0 (109.1-116.9)	140.5 (139.2-141.7)	139.7 (138.2-141.3)
113	113.1 (109.4-116.8)	113.1 (109.1-117.1)	140.5 (139.2-141.8)	139.8 (138.2-141.4)
114	113.2 (109.3-117.0)	113.2 (109.0-117.3)	140.6 (139.3-142.0)	139.9 (138.2-141.5)
115	113.2 (109.3-117.2)	113.2 (109.0-117.5)	140.7 (139.3-142.1)	139.9 (138.3-141.6)
116	113.3 (109.3-117.4)	113.3 (108.9-117.7)	140.8 (139.4-142.2)	140.0 (138.3-141.7)
117	113.4 (109.2-117.6)	113.4 (108.9-117.9)	140.8 (139.4-142.3)	140.1 (138.3-141.8)
118	113.5 (109.2-117.8)	113.5 (108.8-118.2)	140.9 (139.5-142.4)	140.1 (138.3-141.9)
119	113.6 (109.1-118.0)	113.6 (108.8-118.4)	141.0 (139.5-142.5)	140.2 (138.4-142.0)
120	113.6 (109.1-118.2)	113.7 (108.7-118.6)	141.1 (139.6-142.6)	140.2 (138.4-142.1)

121	113.7 (109.0-118.4)	113.7 (108.7-118.8)	141.2 (139.6-142.7)	140.3 (138.4-142.2)
122	113.8 (109.0-118.6)	113.8 (108.6-119.0)	141.2 (139.7-142.8)	140.4 (138.4-142.3)
123	113.9 (108.9-118.8)	113.9 (108.5-119.3)	141.3 (139.7-142.9)	140.4 (138.5-142.4)
124	114.0 (108.9-119.0)	114.0 (108.5-119.5)	141.4 (139.7-143.0)	140.5 (138.5-142.5)
125	114.0 (108.8-119.2)	114.1 (108.4-119.7)	141.5 (139.8-143.1)	140.6 (138.5-142.6)
126	114.1 (108.8-119.4)	114.1 (108.4-119.9)	141.5 (139.8-143.2)	140.6 (138.5-142.7)
127	114.2 (108.7-119.6)	114.2 (108.3-120.2)	141.6 (139.9-143.3)	140.7 (138.5-142.8)
128	114.3 (108.7-119.8)	114.3 (108.2-120.4)	141.7 (139.9-143.4)	140.7 (138.6-142.9)
129	114.3 (108.6-120.1)	114.4 (108.2-120.6)	141.8 (139.9-143.6)	140.8 (138.6-143.0)
130	114.4 (108.6-120.3)	114.5 (108.1-120.8)	141.8 (140.0-143.7)	140.9 (138.6-143.1)
131	114.5 (108.5-120.5)	114.6 (108.1-121.1)	141.9 (140.0-143.8)	140.9 (138.6-143.2)
Females				
0	98.7 (96.0-101.4)	97.6 (95.0-100.2)	132.2 (130.6-133.8)	132.8 (130.3-135.3)
1	98.9 (96.3-101.6)	97.9 (95.3-100.4)	132.3 (130.8-133.9)	132.9 (130.5-135.4)
2	99.1 (96.6-101.7)	98.1 (95.7-100.6)	132.5 (130.9-134.0)	133.0 (130.6-135.4)
3	99.4 (96.9-101.8)	98.4 (96.0-100.8)	132.6 (131.1-134.1)	133.1 (130.8-135.4)
4	99.6 (97.2-102.0)	98.6 (96.3-100.9)	132.7 (131.3-134.2)	133.2 (131.0-135.4)
5	99.8 (97.5-102.1)	98.9 (96.7-101.1)	132.9 (131.5-134.3)	133.3 (131.1-135.5)
6	100.0 (97.8-102.3)	99.1 (97.0-101.3)	133.0 (131.7-134.4)	133.4 (131.3-135.5)
7	100.2 (98.1-102.4)	99.4 (97.3-101.5)	133.1 (131.8-134.4)	133.5 (131.5-135.5)
8	100.5 (98.4-102.5)	99.7 (97.6-101.7)	133.3 (132.0-134.5)	133.6 (131.6-135.5)
9	100.7 (98.7-102.7)	99.9 (98.0-101.8)	133.4 (132.2-134.6)	133.7 (131.8-135.6)
10	100.9 (99.0-102.8)	100.2 (98.3-102.0)	133.5 (132.4-134.7)	133.8 (132.0-135.6)
11	101.1 (99.3-103.0)	100.4 (98.6-102.2)	133.7 (132.6-134.8)	133.9 (132.1-135.6)
12	101.3 (99.6-103.1)	100.7 (99.0-102.4)	133.8 (132.7-134.9)	134.0 (132.3-135.7)

13	101.6 (99.9-103.3)	100.9 (99.3-102.6)	133.9 (132.9-135.0)	134.1 (132.4-135.7)
14	101.8 (100.2-103.4)	101.2 (99.6-102.8)	134.1 (133.1-135.1)	134.2 (132.6-135.7)
15	102.0 (100.4-103.6)	101.4 (99.9-103.0)	134.2 (133.3-135.2)	134.3 (132.8-135.8)
16	102.2 (100.7-103.7)	101.7 (100.3-103.1)	134.3 (133.4-135.3)	134.4 (132.9-135.8)
17	102.4 (101.0-103.9)	102.0 (100.6-103.3)	134.5 (133.6-135.3)	134.4 (133.1-135.8)
18	102.7 (101.3-104.0)	102.2 (100.9-103.5)	134.6 (133.8-135.4)	134.5 (133.2-135.9)
19	102.9 (101.6-104.2)	102.5 (101.2-103.7)	134.7 (133.9-135.5)	134.6 (133.4-135.9)
20	103.1 (101.8-104.3)	102.7 (101.5-103.9)	134.9 (134.1-135.6)	134.7 (133.5-135.9)
21	103.3 (102.1-104.5)	103.0 (101.8-104.1)	135.0 (134.3-135.7)	134.8 (133.7-136.0)
22	103.5 (102.4-104.7)	103.2 (102.1-104.3)	135.1 (134.4-135.8)	134.9 (133.8-136.0)
23	103.7 (102.6-104.8)	103.4 (102.4-104.5)	135.3 (134.6-135.9)	135.0 (133.9-136.1)
24	103.9 (102.9-105.0)	103.7 (102.6-104.7)	135.4 (134.7-136)	135.1 (134.1-136.1)
25	104.1 (103.1-105.2)	103.9 (102.9-104.9)	135.5 (134.9-136.1)	135.2 (134.2-136.2)
26	104.3 (103.3-105.3)	104.2 (103.2-105.1)	135.6 (135-136.2)	135.3 (134.3-136.2)
27	104.5 (103.5-105.5)	104.4 (103.4-105.3)	135.7 (135.1-136.3)	135.4 (134.5-136.3)
28	104.7 (103.8-105.7)	104.6 (103.7-105.5)	135.9 (135.3-136.5)	135.5 (134.6-136.4)
29	104.9 (104.0-105.9)	104.8 (103.9-105.8)	136.0 (135.4-136.6)	135.6 (134.7-136.4)
30	105.1 (104.2-106.0)	105.1 (104.1-106.0)	136.1 (135.5-136.7)	135.6 (134.8-136.5)
31	105.3 (104.3-106.2)	105.3 (104.4-106.2)	136.2 (135.6-136.8)	135.7 (134.9-136.6)
32	105.5 (104.5-106.4)	105.5 (104.6-106.4)	136.3 (135.8-136.9)	135.8 (135.0-136.6)
33	105.6 (104.7-106.6)	105.7 (104.8-106.6)	136.4 (135.9-137.0)	135.9 (135.1-136.7)
34	105.8 (104.9-106.8)	105.9 (105.0-106.8)	136.5 (136.0-137.1)	136.0 (135.2-136.8)
35	106.0 (105.0-106.9)	106.1 (105.2-107.0)	136.6 (136.1-137.2)	136.1 (135.3-136.9)
36	106.1 (105.2-107.1)	106.3 (105.4-107.2)	136.7 (136.2-137.3)	136.1 (135.3-137.0)
37	106.3 (105.3-107.3)	106.5 (105.5-107.4)	136.8 (136.3-137.4)	136.2 (135.4-137.0)

38	106.4 (105.5-107.4)	106.6 (105.7-107.6)	136.9 (136.3-137.5)	136.3 (135.5-137.1)
39	106.6 (105.6-107.6)	106.8 (105.9-107.8)	137.0 (136.4-137.6)	136.4 (135.6-137.2)
40	106.7 (105.7-107.7)	107.0 (106.0-107.9)	137.1 (136.5-137.7)	136.5 (135.6-137.3)
41	106.9 (105.9-107.9)	107.2 (106.2-108.1)	137.2 (136.6-137.8)	136.5 (135.7-137.4)
42	107.0 (106.0-108.0)	107.3 (106.3-108.3)	137.3 (136.7-137.9)	136.6 (135.8-137.4)
43	107.1 (106.1-108.2)	107.5 (106.5-108.4)	137.4 (136.7-138.0)	136.7 (135.8-137.5)
44	107.2 (106.2-108.3)	107.6 (106.6-108.6)	137.4 (136.8-138.0)	136.7 (135.9-137.6)
45	107.4 (106.3-108.4)	107.8 (106.8-108.7)	137.5 (136.9-138.1)	136.8 (136.0-137.7)
46	107.5 (106.4-108.5)	107.9 (106.9-108.9)	137.6 (137.0-138.2)	136.9 (136.0-137.7)
47	107.6 (106.5-108.6)	108.0 (107.0-109.0)	137.6 (137.0-138.3)	136.9 (136.1-137.8)
48	107.7 (106.6-108.7)	108.2 (107.2-109.2)	137.7 (137.1-138.3)	137.0 (136.1-137.9)
49	107.8 (106.7-108.8)	108.3 (107.3-109.3)	137.8 (137.1-138.4)	137.1 (136.2-137.9)
50	107.9 (106.8-108.9)	108.4 (107.4-109.4)	137.8 (137.2-138.5)	137.1 (136.3-138.0)
51	108.0 (106.9-109.0)	108.5 (107.5-109.5)	137.9 (137.3-138.5)	137.2 (136.3-138.1)
52	108.1 (107.0-109.1)	108.6 (107.6-109.6)	137.9 (137.3-138.6)	137.3 (136.4-138.1)
53	108.1 (107.1-109.2)	108.7 (107.7-109.7)	138.0 (137.4-138.6)	137.3 (136.4-138.2)
54	108.2 (107.2-109.3)	108.8 (107.8-109.8)	138.1 (137.4-138.7)	137.4 (136.5-138.2)
55	108.3 (107.2-109.3)	108.9 (107.9-109.9)	138.1 (137.5-138.7)	137.4 (136.5-138.3)
56	108.4 (107.3-109.4)	109.0 (108.0-110.0)	138.1 (137.5-138.8)	137.5 (136.6-138.4)
57	108.4 (107.4-109.5)	109.1 (108.1-110.1)	138.2 (137.6-138.8)	137.5 (136.7-138.4)
58	108.5 (107.4-109.5)	109.2 (108.2-110.2)	138.2 (137.6-138.9)	137.6 (136.7-138.5)
59	108.6 (107.5-109.6)	109.3 (108.3-110.3)	138.3 (137.7-138.9)	137.6 (136.8-138.5)
60	108.6 (107.6-109.7)	109.4 (108.3-110.4)	138.3 (137.7-138.9)	137.7 (136.8-138.6)
61	108.7 (107.6-109.7)	109.4 (108.4-110.5)	138.4 (137.7-139.0)	137.7 (136.9-138.6)
62	108.7 (107.6-109.8)	109.5 (108.5-110.5)	138.4 (137.8-139.0)	137.8 (136.9-138.7)

63	108.8 (107.7-109.9)	109.6 (108.5-110.6)	138.4 (137.8-139.1)	137.8 (137.0-138.7)
64	108.8 (107.7-109.9)	109.6 (108.6-110.7)	138.5 (137.8-139.1)	137.9 (137.0-138.7)
65	108.9 (107.7-110.0)	109.7 (108.6-110.8)	138.5 (137.9-139.1)	137.9 (137.1-138.8)
66	108.9 (107.8-110.0)	109.8 (108.7-110.8)	138.5 (137.9-139.2)	138.0 (137.2-138.8)
67	108.9 (107.8-110.1)	109.8 (108.7-110.9)	138.5 (137.9-139.2)	138.0 (137.2-138.9)
68	109.0 (107.8-110.1)	109.9 (108.8-111.0)	138.6 (137.9-139.2)	138.1 (137.3-138.9)
69	109.0 (107.8-110.2)	109.9 (108.8-111.0)	138.6 (137.9-139.3)	138.1 (137.3-139.0)
70	109.0 (107.8-110.3)	110.0 (108.8-111.1)	138.6 (138.0-139.3)	138.2 (137.3-139.0)
71	109.1 (107.8-110.3)	110.0 (108.9-111.2)	138.6 (138.0-139.3)	138.2 (137.4-139.1)
72	109.1 (107.8-110.4)	110.1 (108.9-111.2)	138.7 (138.0-139.4)	138.3 (137.4-139.1)
73	109.1 (107.8-110.4)	110.1 (108.9-111.3)	138.7 (138.0-139.4)	138.3 (137.5-139.1)
74	109.1 (107.7-110.5)	110.1 (108.9-111.4)	138.7 (138.0-139.4)	138.4 (137.5-139.2)
75	109.1 (107.7-110.6)	110.2 (108.9-111.4)	138.7 (138.0-139.5)	138.4 (137.5-139.2)
76	109.2 (107.7-110.6)	110.2 (108.9-111.5)	138.7 (138.0-139.5)	138.4 (137.6-139.3)
77	109.2 (107.7-110.7)	110.2 (108.9-111.6)	138.8 (138.0-139.5)	138.5 (137.6-139.3)
78	109.2 (107.6-110.8)	110.3 (108.9-111.7)	138.8 (138.0-139.6)	138.5 (137.6-139.4)
79	109.2 (107.6-110.8)	110.3 (108.9-111.7)	138.8 (138.0-139.6)	138.6 (137.7-139.4)
80	109.2 (107.5-110.9)	110.3 (108.9-111.8)	138.8 (138.0-139.6)	138.6 (137.7-139.5)
81	109.2 (107.5-111.0)	110.4 (108.9-111.9)	138.8 (137.9-139.7)	138.6 (137.7-139.5)
82	109.2 (107.4-111.1)	110.4 (108.8-112.0)	138.8 (137.9-139.7)	138.7 (137.8-139.6)
83	109.2 (107.4-111.1)	110.4 (108.8-112.0)	138.8 (137.9-139.7)	138.7 (137.8-139.7)
84	109.3 (107.3-111.2)	110.4 (108.8-112.1)	138.8 (137.9-139.8)	138.8 (137.8-139.7)
85	109.3 (107.3-111.3)	110.5 (108.7-112.2)	138.8 (137.9-139.8)	138.8 (137.8-139.8)
86	109.3 (107.2-111.3)	110.5 (108.7-112.3)	138.9 (137.8-139.9)	138.8 (137.8-139.8)
87	109.3 (107.1-111.4)	110.5 (108.7-112.3)	138.9 (137.8-139.9)	138.9 (137.8-139.9)

88	109.3 (107.1-111.5)	110.5 (108.6-112.4)	138.9 (137.8-140.0)	138.9 (137.9-140.0)
89	109.3 (107.0-111.6)	110.5 (108.6-112.5)	138.9 (137.8-140.0)	138.9 (137.9-140.0)
90	109.3 (106.9-111.6)	110.6 (108.6-112.6)	138.9 (137.7-140.0)	139.0 (137.9-140.1)
91	109.3 (106.8-111.7)	110.6 (108.5-112.6)	138.9 (137.7-140.1)	139.0 (137.9-140.2)
92	109.3 (106.8-111.8)	110.6 (108.5-112.7)	138.9 (137.7-140.1)	139.1 (137.9-140.2)
93	109.3 (106.7-111.9)	110.6 (108.4-112.8)	138.9 (137.7-140.2)	139.1 (137.9-140.3)
94	109.3 (106.6-112.0)	110.6 (108.4-112.9)	138.9 (137.6-140.2)	139.1 (137.9-140.4)
95	109.3 (106.5-112.0)	110.7 (108.3-113.0)	138.9 (137.6-140.3)	139.2 (137.9-140.4)
96	109.3 (106.5-112.1)	110.7 (108.3-113.1)	138.9 (137.6-140.3)	139.2 (137.9-140.5)
97	109.3 (106.4-112.2)	110.7 (108.2-113.1)	138.9 (137.5-140.3)	139.2 (137.9-140.6)
98	109.3 (106.3-112.3)	110.7 (108.2-113.2)	138.9 (137.5-140.4)	139.3 (137.9-140.7)
99	109.3 (106.2-112.4)	110.7 (108.1-113.3)	139.0 (137.5-140.4)	139.3 (137.9-140.7)
100	109.3 (106.2-112.4)	110.7 (108.1-113.4)	139.0 (137.5-140.5)	139.4 (137.9-140.8)
101	109.3 (106.1-112.5)	110.8 (108.0-113.5)	139.0 (137.4-140.5)	139.4 (137.9-140.9)
102	109.3 (106.0-112.6)	110.8 (108.0-113.6)	139.0 (137.4-140.6)	139.4 (137.9-141.0)
103	109.3 (105.9-112.7)	110.8 (108.0-113.6)	139.0 (137.4-140.6)	139.5 (137.9-141.0)
104	109.3 (105.8-112.8)	110.8 (107.9-113.7)	139.0 (137.3-140.7)	139.5 (137.9-141.1)
105	109.3 (105.8-112.8)	110.8 (107.9-113.8)	139.0 (137.3-140.7)	139.5 (137.9-141.2)
106	109.3 (105.7-112.9)	110.8 (107.8-113.9)	139.0 (137.3-140.8)	139.6 (137.9-141.3)
107	109.3 (105.6-113.0)	110.9 (107.8-114.0)	139.0 (137.2-140.8)	139.6 (137.9-141.3)
108	109.3 (105.5-113.1)	110.9 (107.7-114.1)	139.0 (137.2-140.9)	139.7 (137.9-141.4)
109	109.3 (105.4-113.2)	110.9 (107.6-114.2)	139.0 (137.2-140.9)	139.7 (137.9-141.5)
110	109.3 (105.4-113.3)	110.9 (107.6-114.2)	139.0 (137.1-140.9)	139.7 (137.9-141.6)
111	109.3 (105.3-113.3)	110.9 (107.5-114.3)	139.0 (137.1-141.0)	139.8 (137.9-141.6)
112	109.3 (105.2-113.4)	111.0 (107.5-114.4)	139.1 (137.1-141.0)	139.8 (137.9-141.7)

113	109.3 (105.1-113.5)	111.0 (107.4-114.5)	139.1 (137.0-141.1)	139.8 (137.9-141.8)
114	109.3 (105.0-113.6)	111.0 (107.4-114.6)	139.1 (137.0-141.1)	139.9 (137.9-141.9)
115	109.3 (105.0-113.7)	111.0 (107.3-114.7)	139.1 (137.0-141.2)	139.9 (137.9-142.0)
116	109.3 (104.9-113.8)	111.0 (107.3-114.8)	139.1 (136.9-141.2)	140.0 (137.9-142.0)
117	109.3 (104.8-113.8)	111.0 (107.2-114.9)	139.1 (136.9-141.3)	140.0 (137.9-142.1)
118	109.3 (104.7-113.9)	111.1 (107.2-114.9)	139.1 (136.9-141.3)	140.0 (137.9-142.2)
119	109.3 (104.6-114)	111.1 (107.1-115)	139.1 (136.8-141.4)	140.1 (137.9-142.3)
120	109.3 (104.6-114.1)	111.1 (107.1-115.1)	139.1 (136.8-141.4)	140.1 (137.9-142.3)
121	109.3 (104.5-114.2)	111.1 (107.0-115.2)	139.1 (136.8-141.5)	140.1 (137.9-142.4)
122	109.3 (104.4-114.3)	111.1 (107.0-115.3)	139.1 (136.7-141.5)	140.2 (137.9-142.5)
123	109.3 (104.3-114.4)	111.2 (106.9-115.4)	139.1 (136.7-141.6)	140.2 (137.9-142.6)
124	109.3 (104.2-114.4)	111.2 (106.9-115.5)	139.1 (136.6-141.6)	140.3 (137.8-142.7)
125	109.3 (104.1-114.5)	111.2 (106.8-115.6)	139.1 (136.6-141.7)	140.3 (137.8-142.7)
126	109.3 (104.1-114.6)	111.2 (106.8-115.7)	139.2 (136.6-141.7)	140.3 (137.8-142.8)
127	109.3 (104.0-114.7)	111.2 (106.7-115.7)	139.2 (136.5-141.8)	140.4 (137.8-142.9)
128	109.3 (103.9-114.8)	111.2 (106.7-115.8)	139.2 (136.5-141.8)	140.4 (137.8-143.0)
129	109.3 (103.8-114.9)	111.3 (106.6-115.9)	139.2 (136.5-141.9)	140.4 (137.8-143.1)
130	109.3 (103.7-115.0)	111.3 (106.5-116.0)	139.2 (136.4-141.9)	140.5 (137.8-143.2)
131	109.3 (103.6-115.0)	111.3 (106.5-116.1)	139.2 (136.4-142.0)	140.5 (137.8-143.2)

Supplemental Table S2. Number of children used in estimations of reference and optimal hemoglobin curves by sex and standardized age groups.

	Reference Curves		**Optimal Curves**	
Age Group, n (%)	**Female N = 2,146**	**Male N = 2,451**	**Female N = 1,628**	**Male N = 1,798**
2 weeks - 2 months	12 (0.6)	21 (0.9)	10 (0.6)	13 (0.7)
3 - 5 months	23 (1.1)	26 (1.1)	14 (0.9)	12 (0.7)
6 - 7 months	67 (3.1)	65 (2.7)	43 (2.6)	38 (2.1)
9 - 11 months	110 (5.1)	138 (5.6)	71 (4.4)	78 (4.3)
12 - 14 months	149 (6.9)	161 (6.6)	93 (5.7)	93 (5.2)
15 - 17 months	107 (5.0)	121 (4.9)	61 (3.7)	74 (4.1)
18 - 20 months	137 (6.4)	160 (6.5)	83 (5.1)	94 (5.2)
21 - 23 months	40 (1.9)	62 (2.5)	22 (1.4)	34 (1.9)
24 - 26 months	189 (8.8)	189 (7.7)	150 (9.2)	147 (8.2)
27 - 29 months	44 (2.1)	49 (2.0)	32 (2.0)	36 (2.0)
30 - 32 months	26 (1.2)	41 (1.7)	19 (1.2)	24 (1.3)
33 - 35 months	25 (1.2)	39 (1.6)	21 (1.3)	27 (1.5)
3 - <4 years	294 (14)	298 (12)	242 (15)	239 (13)
4 - <5 years	259 (12)	322 (13)	211 (13)	267 (15)
5 - <6 years	228 (11)	252 (10)	172 (11)	197 (11)
6 - <7 years	113 (5.3)	143 (5.8)	93 (5.7)	119 (6.6)
7 - <8 years	114 (5.3)	102 (4.2)	104 (6.4)	86 (4.8)
8 - <9 years	100 (4.7)	107 (4.4)	89 (5.5)	87 (4.8)
9 - <10 years	65 (3.0)	88 (3.6)	58 (3.6)	75 (4.2)
10 - <11 years	44 (2.1)	67 (2.7)	40 (2.5)	58 (3.2)

CHAPTER 6: ESTABLISHING FERRITIN REFERENCE AND OPTIMAL CURVES IN PRE-ADOLESCENTS: PROVIDING CONTEXT TO CLINICAL DECISION THRESHOLDS

6.1. Preface to Chapter 6

Chapter 6 presents the fifth manuscript of this thesis, addressing the second objective of our research. In this manuscript, we establish reference curves for ferritin based on a cohort of healthy children in Canada. In addition to conventional reference intervals and curves, we develop optimal reference curves. In addition, we created an interactive Shiny web application that allows a user to obtain both traditional and optimal ferritin intervals for a child of any age between 2-week to 10.99 years old (Optimized Reference Assessment for Clinical Laboratory Evaluation of Ferritin [ORACLE-FER]: https://svetrcocfer.shinyapps.io/ORACLE-FER/). The web tool also supports a multi measure assessment approach, allowing individual laboratory results to be evaluated simultaneously against traditional reference intervals, optimal intervals, and global WHO thresholds. This Manuscript is currently submitted to *Blood*.

VB led the study design, review and selection of optimality criteria, performed all analysis including estimation of reference and optimal interval and curves and missing data imputation, development of Shiny app, presentation and interpretation of results and has drafted the initial manuscript. BKP, PCP, and JSH were involved in the design, data interpretation, writing and revision of the manuscript. FM and ML were involved in the design, data interpretation and revision of the manuscript. CMB, CSB, and JLM were involved in acquisition of data and revision of the manuscript.

6.2 Manuscript status: revision submitted to *JAMA Network Open*

Ferritin reference and optimal curves in pre-adolescents: contextualizing iron deficiency thresholds

Vid Bijelić PhD[1,2,4], Franco Momoli PhD[1], Mira Liebman MD[2], Beth K. Potter PhD[1], Cornelia M. Borkhoff PhD[4], Catherine S. Birken MD MSc[4], Jonathon L. Maguire MD MSc[5], Patricia C. Parkin MD[4,*], Jemila S. Hamid PhD[2,3]

1 School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
2 Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada
3 Department of Mathematics and Statistics, University of Ottawa, Ottawa, Ontario, Canada
4 The Hospital for Sick Children Research Institute, Toronto, Ontario, Canada
5 MAP Centre for Urban Health Solutions, St. Michael's Hospital, Unity Health Toronto

* Corresponding Author:
Patricia C. Parkin, MD, FRCPC
The Hospital for Sick Children Research Institute
686 Bay St, Toronto, ON, M5G 0A4, Canada
416-813-7654
Email: patricia.parkin@sickkids.ca

Date of revision: March 13, 2026

Word count: 2,994

KEY POINTS

Question: What are the lower and upper limits of the ferritin reference curves for pre-adolescents and how do these relate to physiologically-based iron deficiency thresholds?

Findings: Limits were highest in early infancy, declined at 1.5 years (lower limit 5-6μg/L), then gradually increased at 9-10 years (lower limit 15-17μg/L). From 3 months to 10 years of age, the lower limit of ferritin was below the physiologically-based threshold of ≤20μg/L.

Meaning: Reference curves, which provide lower and upper limits continuously across age, may improve clinical utility; however, lower limits are below the proposed threshold for identifying iron deficiency in pediatrics.

ABSTRACT

Importance: Ferritin is commonly used to assess iron status. For interpretation, clinicians traditionally rely on age-partitioned reference intervals (RIs) which report 2.5th (lower) and 97.5th (upper) percentiles.

Objectives: To estimate sex-specific continuous ferritin reference curves (RCs) and optimal curves (OCs) for pre-adolescents; to create age- and sex-specific RIs and optimal intervals (OIs); to interpret our estimates alongside ferritin thresholds for iron deficiency; to develop an interactive web-based computational tool.

Design: Cross-sectional analysis using data from a longitudinal cohort study.

Setting: Children attending scheduled primary care health supervision visits from June 2008 to February 2020 in Toronto, Canada.

Participants: 4935 children ages 2-weeks to 10-years.

Exposure: Blood samples were collected and analyzed for ferritin, hemoglobin and C-reactive protein. Parents completed a questionnaire collecting optimality criteria for iron status.

Main Outcomes: Sex-specific RCs were estimated using generalized additive models for location, scale, and shape (GAMLSS). OCs were derived from a sub-sample of participants meeting predefined optimality criteria, using multiple imputation to address missing data. RCs and OCs were interpreted in the context of the WHO threshold (<12μg/L) and the proposed American Society of Hematology (ASH) threshold (≤20μg/L).

Results: RCs were estimated from 4935 children (2322 females, 2613 males) aged 2-weeks to 10-years (median 37 months). OCs were estimated from a subset of 3630 children (1909 males, 1721 females). Curves were highest in early infancy, declined sharply at 1.5 years (lower limit 5-

6µg/L), then gradually increased at 9-10 years (lower limit 15-17µg/L). RCs and OCs showed similar trajectories, with slightly higher limits for OCs. The proportion of pre-adolescents in the optimal population who would be classified as iron deficient for existing (WHO) and proposed (ASH) thresholds differed with age. Using the ASH threshold in this cohort, 55% males and 37% females would be considered iron deficient at 18 months.

Conclusion and Relevance: In this cross-sectional study, we estimated ferritin RCs and OCs from a low-risk pre-adolescent population. The lower limits of curves and intervals were substantially lower than physiologically-based thresholds to define ID. These findings further inform the current discussion regarding distributional versus physiological approaches to interpreting ferritin measurements and guiding clinical decision-making.

INTRODUCTION

Ferritin, measured in serum or plasma, is commonly used in clinical practice to assess iron stores, with low levels indicating iron deficiency (ID) and high levels indicating iron-overload or inflammation.[1,2] In pediatrics, ID prevalence is highest in early childhood for both sexes and for adolescent females. ID is the leading cause of anemia, and is associated with neurodevelopmental impairment, fatigue and decreased concentration.[1] Iron-overload is much less common in children but high levels of ferritin are associated with inflammation, malignancy, and obesity.[1]

To interpret ferritin levels, clinicians rely on reference intervals (RIs) developed following Clinical Laboratory Standards Institute (CLSI) guidelines,[3,4] from samples of healthy reference populations with adequate sample sizes.[3] RIs are defined by the 2.5th and 97.5th percentiles corresponding to lower and upper RI limits, with 90% confidence intervals (CI) for each limit. Sex- and age-specific partitions may be required, particularly in the pediatric population. Our systematic review of pediatric RIs for ferritin identified different age partitions across studies and significant heterogeneity among reported RIs, limiting our ability to perform meta-analysis to generate reliable pooled estimates.[5]

Reference curves (RCs), which model biomarkers as a continuous function of age, offer an alternative to age-partitioned RIs, especially important in pediatrics considering physiological changes from infancy through adolescence.[6,7] The method is similar to the approach taken for developing the World Health Organization (WHO) growth standards.[8] In our systematic review, we identified four studies on pediatric RCs for ferritin and only one provided CIs for inclusion in our meta-analysis.[9]

An alternative approach, described in anthropometric research, involves selecting a sub-sample of individuals free from health conditions and environmental or socioeconomic risk factors related to the biomarker of interest.[10,11] Our systematic review found no studies that excluded children at risk of ID or anemia, suggesting existing estimates may not reflect optimal iron health. Applying optimality criteria may provide more clinically meaningful reference limits.

Emerging evidence suggests that reliance on the distribution-based lower limit of the ferritin RI (2.5th percentile) leads to underdiagnosis of ID in children and adults, with experts advocating for physiologically-based thresholds (Clinical Decision Limits).[12–25] The WHO recommends a ferritin threshold of <12µg/L to define ID for children younger than 5 years and <15µg/L over 5 years, largely based on consensus.[1] The 2025 American Society of Hematology (ASH) draft recommendations suggest a ferritin threshold of ≤20µg/L for children ages 9 months to 4 years, based on a single pediatric study comparing ferritin with bone marrow iron.[26] Using these thresholds rather than the lower RI limit will substantially increase the number of children considered to have ID.[18]

Our objectives were to estimate sex-specific ferritin RCs for pre-adolescents; estimate optimal curves (OCs), using pre-defined optimality criteria; create age- and sex-specific RIs and optimal intervals (OIs); contextualize our estimates alongside WHO and proposed ASH ID thresholds; and develop an interactive web-based computational and graphical tool.

METHODS

Population and study design

We used cross-sectional data from healthy children who participated in longitudinal cohort study from a primary care research network in Toronto, Canada.[27] Children were recruited at any of the following scheduled health supervision visits: 2-weeks, 2-, 4-, 6-, 9-, 12-, 15-, 18-months of age, then annually to 5 years. Parents completed questionnaires covering demographics and medical history, including birthweight, gestational age, and household income. Trained research assistants in participating clinics collected anthropometric measures and blood samples at multiple visits, with blood sample collection being optional. Pre-COVID19 data (June 2008-February 2020) from children 2-weeks to 10-years were included. Children with health conditions affecting growth, acute or chronic conditions (other than asthma and high-functioning autism), or severe developmental delay, and those with families unable to communicate in English were excluded.[27] Blood samples were collected in serum-separator and Ethylenediaminetetraacetic acid (EDTA) tubes, transported at room temperature on the same day to the accredited research laboratory at Mount Sinai Services, Toronto, Canada, and analysed within 4-6 hours for ferritin. Hemoglobin was measured in whole blood using the Sysmex XN-9000 Analyzers (Japan).[27] Serum ferritin and c-reactive protein were measured using a modular platform Roche Diagnostics (Switzerland).[27]

The Research Ethics Boards at the Hospital for Sick Children and St Michael's Hospital granted approval and consent was obtained from parents of participating children. We followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.[28]

Optimal curve sub-group

To create a sub-group at low risk of ID, we used pre-specified optimality criteria obtained from previous research[29–32] and recommendations[33–35] regarding factors associated with iron status. Data were excluded from children with anemia (hemoglobin for ages 6-23 months <105g/L, 24-59 months <110g/L, and ≥60 months <115g/L)[36], living in low-income households (<$42,000 Canadian)[37], and/or under or overweight (above or below zBMI+/-1.96*SD, respectively).[31] BMI (kg/m²) was calculated from measured height/length and weight standardized for age and sex using WHO growth standards.[27,31,38] The income threshold approximates the average Statistics Canada low income cut-off before tax for a family of four residing in urban area with population over 100,000 between 2008-2020.[37] We excluded data from children <2 years if born prematurely (<37 weeks gestational age) or with low birth weight (<2,500grams);[35] or if the C-reactive protein (CRP) was above 5mg/L suggesting acute systemic inflammation which can falsely elevate ferritin.[39]

Data analysis

Some children had repeated measurements, therefore we selected one ferritin measurement per child prioritizing observations from ages with sparse data (referred to as 'lean pick').[5] We estimated sex-specific RCs and OCs using the same analytic methods. RCs included the overall cohort. OCs included the sub-group meeting optimality criteria. The generalized additive models for location, scale and shape (GAMLSS) method was used to model the age-dependent distribution of ferritin separately for males and females. We applied the Box-Cox-t (BCTo) transformation that accommodates skewed and heavy tailed distributions, and selected the log-link function based on generalized Akaike information criteria (AIC).[40] We used the *gamlss* package in R version 5.4-22.[41,42] Parameters of distributions were estimated as a smooth function of age using penalized B-splines.[43] Outliers were identified using Dunn-Smyth residuals

(+/-3 z-scores) and removed.[44] The model was refitted to remaining observations. This process was iteratively repeated until no additional outliers were identified. Ninety percent CIs[3] for the fitted centile curves were estimated using the non-parametric bootstrap method (200 bootstrap samples).

We estimated 5th, 10th and 20th centile curves to allow comparisons with WHO and ASH ID thresholds for children 2-weeks to 5-years.[45] Informed by previous studies,[46,47] we derived age- and sex-specific RIs and OIs from the RCs and OCs by estimating ferritin values at 2.5th and 97.5th centiles using predictions from fitted GAMLSS models. Lower and upper limits were calculated for each month of age and for previously proposed standardized age-intervals[5]: 3-month age intervals for children under 3-years, and one-year intervals afterwards.

Using the Shiny application, we developed a web-based computational and graphical tool to visualize estimated curves, comparison with WHO and ASH thresholds, and allow interactive age- and sex-specific estimation of RIs and OIs based on fitted models. Statistical analyses were performed using R version 4.4.1.[42] RCs were compared descriptively through visual inspection across the age range. No formal statistical tests were performed.

Missing data

RCs were estimated using data from participants with at least one ferritin measurement, maximizing our sample size for younger children using a lean-pick approach. For OCs, we used multiple imputations (MI) by chained equations to handle missing data for variables used to define the optimality criteria.[48,49] In our MI model, we incorporated available repeated measurements and both time-varying and time-invariant variables. Restricted-cubic-splines were used to model non-linear relationships. MI models included age in months and ferritin (both with

knot terms), hemoglobin, household income (defined as the middle of the reporting interval), zBMI score, prematurity, and low birthweight.

To manage missing ferritin data when imputing optimality criteria variables, our primary OC estimation analysis followed the multiple-imputation-then-deletion (MID) approach under the missing-at-random (MAR) assumption.[50] Initially, we relied on the full dataset including participants with and without ferritin values to impute data on optimality criteria. We then excluded participants with no ferritin values or missing CRP values to estimate OCs. Two sensitivity analyses were conducted to assess robustness of the findings to this strategy for managing missing data: deletion-then-multiple-imputation (DMI), where records with missing ferritin values were excluded prior to imputation; and complete-case analysis, where only participants with observed optimality criteria were included in OC estimation. Eighty-one imputed datasets were generated in the primary MID approach and twenty-seven in the DMI approach.[51] For each approach, results were combined using Rubin rules.[48] Imputation was implemented using R *mice* package version 3.17.0, *miceadds* version 3.17.44.[49], and for restricted-cubic-splines *rms* package version 6.8-2.[52]

RESULTS

Study characteristics

Of 11,802 children enrolled, blood work was obtained including ferritin for 4935 children (2322 females, 2613 males) ages 2-weeks to 10-years (median 37 months); we used these values to estimate sex-specific RCs (Table 1). Children in the cohort with and without ferritin appeared to have similar characteristics. Sample size by age is shown in eTable 1.

Sex-specific OCs were estimated using a sub-sample of 3630 children (1909 males, 1721 females) meeting the optimality criteria. From 3.8% (low birthweight) to 9.5% (household

income below low-income cut-off) of participants were in a non-optimal category when each variable was considered separately, with 26% excluded due to being in a non-optimal category for one or more of the criteria (Table 1). Eight males and three females, considered extreme outlier observations, were removed in both RC and OC analyses.

Table 1. Characteristics of study participants.

Characteristics	Full TARGetKids! Cohort[f] N = 11,802	Participants with ferritin values[c] N = 4,935	Participants with ferritin values, imputation[d]	
			All participants N = 4,935	Participants meeting optimality criteria[e] N = 3,630
Age in months, median (IQR)	34 (11, 69)	37 (18, 62)	37 (18, 62)	40 (21, 64)
Sex, n (%)				
Female	5,672 (48.1)	2,322 (47.1)	2,322 (47.1)	1,721 (47.4)
Male	6,130 (51.9)	2,613 (52.9)	2,613 (52.9)	1,909 (52.6)
Ferritin (ug/L), median (IQR)	30 (21, 43)	29 (20, 41)	29 (20, 41)	29 (21, 40)
Missing, n (%)	9,384 (79.5)	0 (0)	0 (0)	0 (0)
CRP (mg/L), median (IQR)	0.3 (0.2, 0.7)	0.3 (0.2, 0.6)	0.3 (0.2, 0.6)	0.2 (0.2, 0.5)
Missing, n (%)	9,121 (77.3)	0 (0)	0 (0)	0 (0)
Low income cut-off (LICO), n (%)				
Above LICO	7,688 (89.1)	3,985 (90.5)	4,464 (90.5)	3,630 (100.0)
Below LICO	941 (10.9)	417 (9.5)	471 (9.5)	0 (0)
Missing	3,173 (26.9)	533 (10.8)	0 (0)	0 (0)
Prematurity[a] n (%)				
Not premature	9,850 (95.9)	4,555 (96.0)	4,734 (95.5)	3,630 (100.0)
Premature	417 (4.1)	189 (4.0)	201 (4.1)	0 (0)
Missing	1,535 (13.0)	191 (3.9)	0 (0)	0 (0)
Low birthweight[a] n (%)				
Not low birthweight	10,285 (96.2)	4,678 (96.3)	4,746 (96.2)	3,630 (100.0)
Low birthweight	409 (3.8)	181 (3.7)	189 (3.8)	0 (0)
Missing	1,108 (9.4)	76 (1.5)	0 (0)	0 (0)
Overweight/Underweight[b] n (%)				
Not Overweight/Underweight	10,313 (90.6)	4,447 (92.0)	4,539 (92.0)	3,630 (100.0)
Overweight/Underweight	1,071 (9.4)	385 (8.0)	396 (8.0)	0 (0)
Missing	418 (3.5)	103 (2.1)	0 (0)	0 (0)
Anemia, n (%)				
No anemia	3,650 (95.6)	4,189 (92.5)	4,569 (92.6)	3,630 (100.0)
Anemia	169 (4.4)	341 (7.5)	366 (7.4)	0 (0)
Missing	7,983 (67.6)	405 (8.2)	0 (0)	0 (0)

[a]Applicable when ferritin measured in first 24 months of life
*[b]Above or Below zBMI +/- 1.96*SD, where zBMI is BMI z-score relative to WHO reference population adjusted for age and sex*
[c]Sample for reference curve estimation
[d]Missing data (not included in denominator for other category) for optimality criteria variables imputed with multiple imputation then deletion (MID) approach (see Methods)
[e]Sample for optimal curve estimation
[f]Cross-sectional sample was derived from the full TARGet Kids! longitudinal dataset using the lean-pick method

Reference and optimal curves

For both males and females aged 2-weeks to 10-years, RCs and OCs showed a similar pattern with respect to age, with higher variation in older females (Figure 1). Although the OC lower limits were consistently above the RC lower limits, the two curves generally overlapped. Both lower and upper limits were high in early infancy and sharply declined during the first year of life, reaching a minimum at approximately 1-1.5 years of age, with the lower limit reaching approximately 5-6µg/L. Both curves gradually increased after the second year of life, reaching a maximum at approximately 9-10 years, with the lower limit reaching approximately 15-17µg/L.

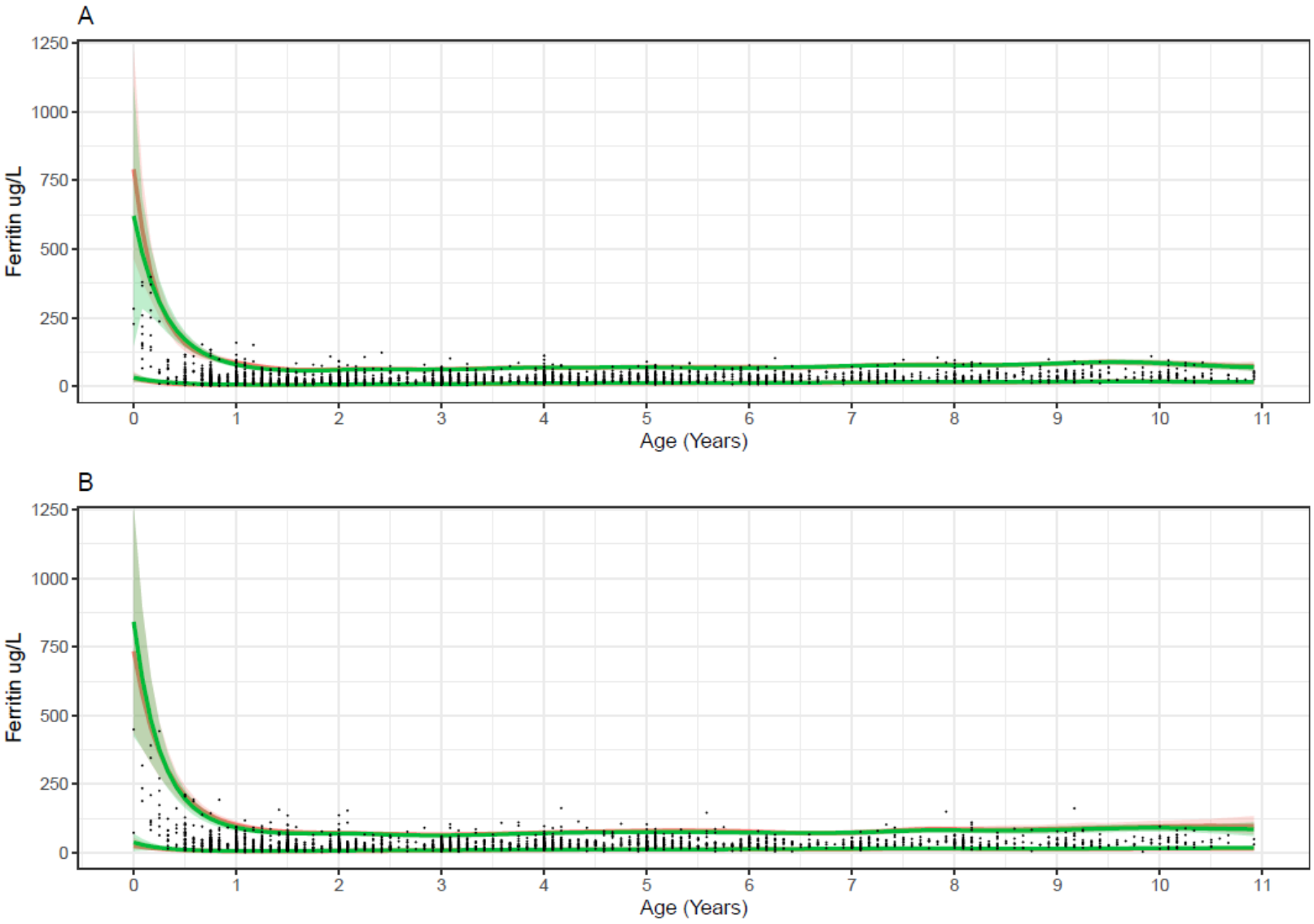


Figure 1. Ferritin reference curves (red) and optimal curves (green) with the corresponding 90% confidence intervals for children aged 2 weeks to 10 years. The actual measurements of ferritin from the 2322 children are presented using scatter plots (black) for 2613 males (A), and 2322 females (B).

ID thresholds

Age-specific 5th, 10th, and 20th percentile thresholds as well as WHO and ASH thresholds for the optimal populations of children 2-weeks to 5-years are shown in Figure 2 (males) and Figure 3 (females). For both males and females, the WHO threshold generally fell between the 5th and 10th percentile of our OCs, but with fluctuation over age. The ASH threshold exceeded

the 20th percentile between 9-months and 3.5 to 4-years of age. The proportion of children in the optimal population who would be classified with ID for each threshold differed by child age. For example, at 18-months of age, the proportion for males and females would be 19.7% and 15.1% using the WHO threshold, and 54.7% and 37.0% using the ASH threshold, respectively.

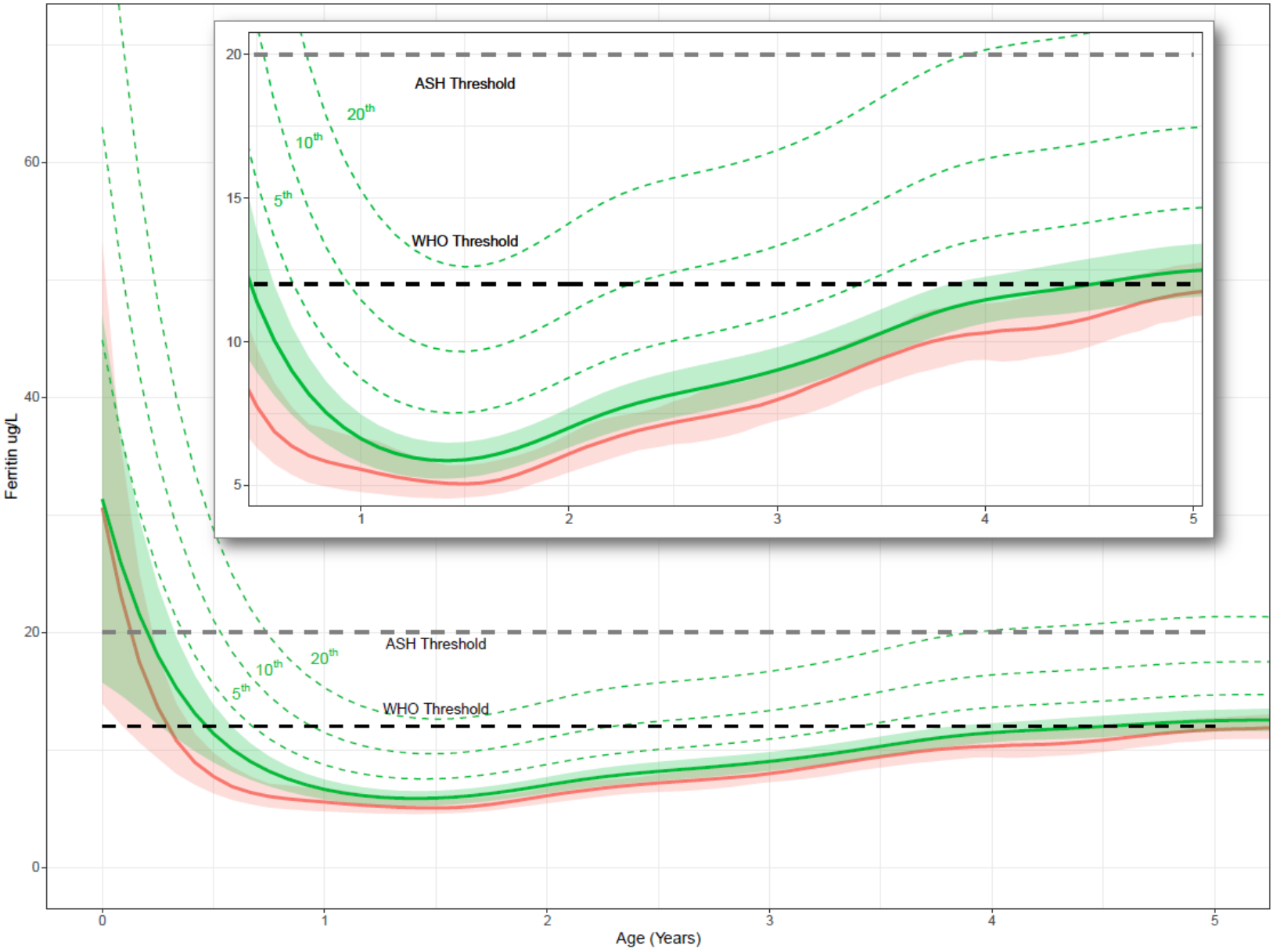


Figure 2. Lower limits of ferritin reference curves (red) and optimal curves (green) for males with the corresponding 90% confidence intervals (shaded) for children aged 2 weeks to 5 years. The WHO and ASH ferritin thresholds for iron deficiency are also presented. Regions between the age of 6 months and 5 years are magnified. Additional optimality centile cures (5th, 10th, and 20th) are shown in green.

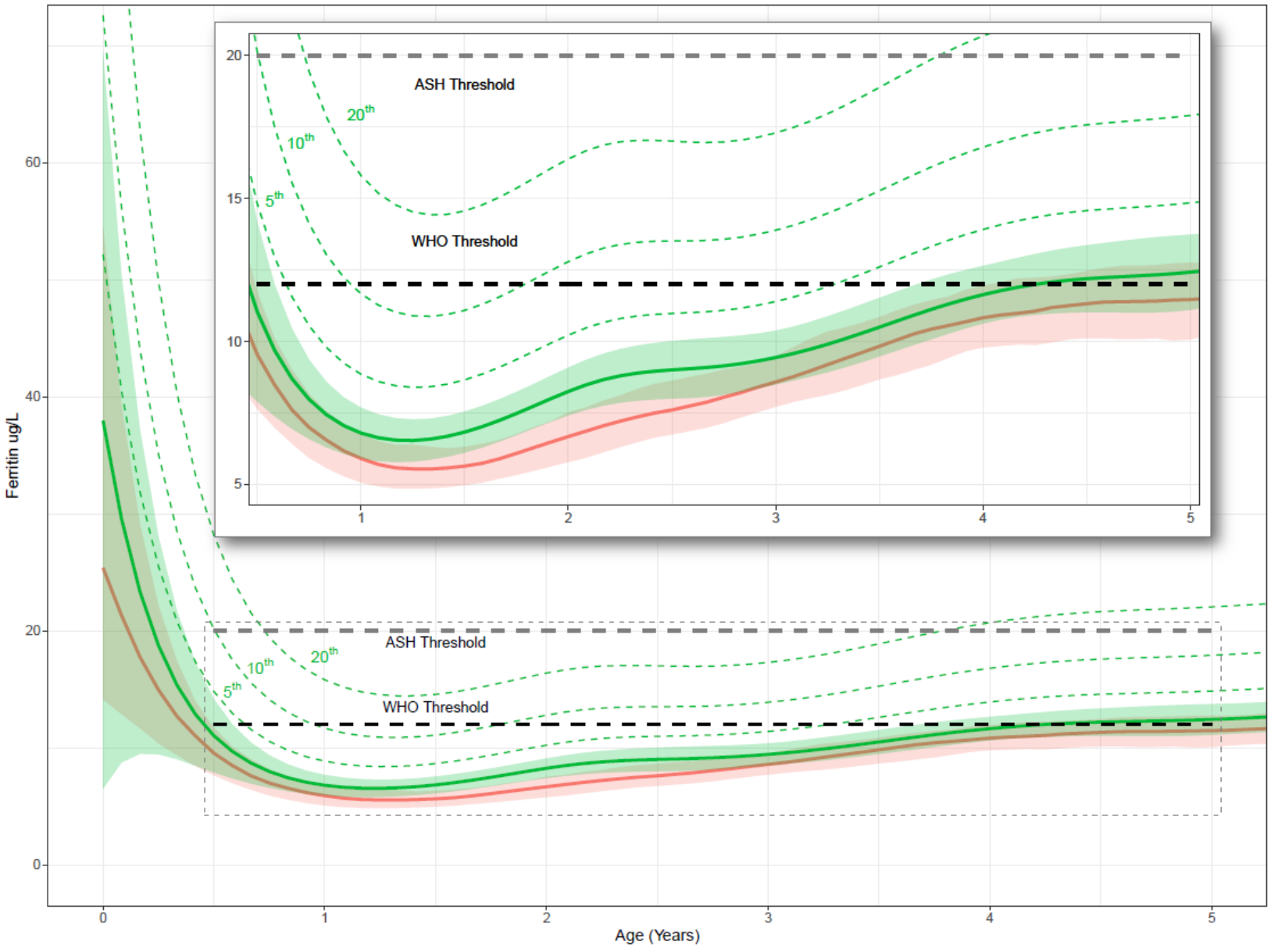


Figure 3. Lower limits of ferritin reference curves (red) and optimal curves (green) for females with the corresponding 90% confidence intervals (shaded) for children aged 2 weeks to 5 years. The WHO and ASH ferritin thresholds for iron deficiency are also presented. Regions between the age of 6 months and 5 years are magnified. Additional optimality centile cures (5^{th}, 10^{th}, and 20^{th}) are shown in green.

Reference intervals

Table 2 provides RIs and OIs (lower and upper limits) with the corresponding 90%CIs for males and females, based on three-month age intervals for children under 3-years and yearly intervals after 3-years. Results using monthly age intervals are presented in eTable 2.

Table 2. Age- and sex-specific reference and optimal intervals for ferritin (µg/L) with the corresponding 90% confidence intervals.

Age Group	Lower		Upper	
	Optimal Limit	Reference Limit	Optimal Limit	Reference Limit
Males				
2 weeks - 2 months	26.2 (14.6-37.9)	23.7 (12.3-38)	496 (231.1-760.9)	592 (389-836)
3 - 5 months	15.5 (11-19.9)	11.1 (8.1-14.4)	252.9 (195.4-310.4)	245.7 (203.4-298.4)
6 - 8 months	10.1 (8.2-12.1)	7 (5.8-8.7)	146.4 (124.2-168.5)	135.7 (113-164.2)
9 - 11 months	7.6 (6.5-8.7)	5.8 (5-6.9)	97.5 (86.1-109)	99 (88.8-112.9)
12 - 14 months	6.3 (5.6-7.1)	5.4 (4.7-6.5)	73.8 (66.5-81.1)	81.3 (73.9-90.6)
15 - 17 months	5.9 (5.2-6.5)	5.1 (4.6-5.8)	62.4 (56.9-67.9)	68.5 (63.3-75)
18 - 20 months	6 (5.4-6.6)	5.1 (4.6-5.8)	58.1 (53.3-63)	61.4 (56.5-67.2)
21 - 23 months	6.5 (5.9-7.1)	5.6 (5-6.2)	58.5 (53.9-63.2)	61 (56.9-66.3)
24 - 26 months	7.2 (6.5-7.9)	6.3 (5.6-7)	60.7 (55.7-65.7)	63.3 (58.9-69.3)
27 - 29 months	7.9 (7.1-8.6)	6.9 (6.2-7.5)	61.8 (56.5-67.2)	63.9 (58.6-70.3)
30 - 32 months	8.3 (7.5-9.1)	7.3 (6.5-8.1)	61.6 (56.2-67)	62.9 (56.9-69.3)
33 - 35 months	8.7 (7.9-9.5)	7.6 (6.9-8.6)	61.4 (56.4-66.4)	62.4 (57.1-68.9)
3 - <4 years	10.2 (9.4-11)	9.2 (8.3-10.2)	65.1 (60.1-70)	66.1 (60.8-72.7)
4 - <5 years	11.9 (11.1-12.8)	10.9 (9.9-11.9)	69.3 (64.6-73.9)	69 (63.8-75.1)
5 - <6 years	12.6 (11.6-13.6)	12 (11-13.2)	68.3 (63.5-73.1)	71.3 (65.8-76.8)
6 - <7 years	13.7 (12.4-15)	12.6 (11.2-14)	70.5 (64.8-76.1)	70.5 (64.3-78.4)
7 - <8 years	15.3 (13.9-16.8)	14.5 (13-16.3)	76.8 (70-83.7)	77.8 (69.8-87.3)
8 - <9 years	15.8 (14.2-17.4)	15.2 (13.3-17)	78.8 (71.5-86.2)	78.7 (70.6-88.2)
9 - <10 years	17.9 (15.9-19.9)	17.3 (15-20.1)	86.9 (78-95.9)	86.9 (76.6-96.9)
10 - <11 years	16.6 (14.1-19.1)	16.5 (13.9-19.4)	76.7 (65.5-88)	78.6 (64-92.2)
Females				
2 weeks - 2 months	30.3 (8.2-52.3)	21.5 (12.9-41)	653.7 (378.2-929.3)	589.2 (382.1-928.3)
3 - 5 months	15.6 (9-22.3)	12.9 (9.4-17.8)	303.8 (237-370.5)	305.6 (239.9-380.2)
6 - 8 months	9.8 (7.4-12.2)	8.5 (7-10.3)	167.5 (140.8-194.2)	180.4 (153.1-208)
9 - 11 months	7.5 (6.3-8.7)	6.6 (5.6-7.6)	111.3 (95.5-127.2)	122.7 (107.6-138.2)
12 - 14 months	6.7 (5.8-7.5)	5.7 (4.9-6.5)	86.3 (75.9-96.7)	95.3 (85.9-106.2)
15 - 17 months	6.6 (5.9-7.3)	5.5 (4.9-6.3)	75.2 (67.5-82.9)	81.7 (74.8-89.7)
18 - 20 months	7 (6.3-7.7)	5.7 (5.1-6.4)	71.1 (64.3-77.9)	75.9 (68.8-82.1)
21 - 23 months	7.7 (7-8.5)	6.3 (5.5-7)	70.4 (63.8-76.9)	73.6 (67.2-79.6)
24 - 26 months	8.5 (7.6-9.3)	6.9 (5.9-7.7)	70.1 (63.3-76.9)	72.3 (65.8-78.9)
27 - 29 months	8.9 (7.9-9.8)	7.3 (6.5-8.3)	68.2 (61.4-75)	70.1 (62.9-77.2)
30 - 32 months	9 (8-10.1)	7.7 (6.9-8.7)	65.3 (58.8-71.9)	68.2 (60.2-75.1)
33 - 35 months	9.2 (8.2-10.2)	8.2 (7.3-9.1)	63.6 (57.8-69.5)	66.9 (59.5-73.3)
3 - <4 years	10.4 (9.4-11.4)	9.7 (8.6-10.7)	66.6 (60.9-72.3)	69.6 (63.2-75.8)
4 - <5 years	12.1 (10.9-13.3)	11.2 (10-12.4)	73.9 (67.9-79.9)	77.3 (70.9-84.6)
5 - <6 years	12.8 (11.5-14.2)	11.8 (10.4-13.1)	75.3 (68.1-82.6)	80.1 (72.6-89.9)
6 - <7 years	14.1 (12.3-15.9)	12.4 (10.9-14.7)	72.6 (65.5-79.7)	73.8 (65.4-82.2)
7 - <8 years	15.8 (13.9-17.6)	14.6 (12.6-17.2)	78.8 (70-87.5)	81.3 (71.8-91.1)
8 - <9 years	14.7 (12.8-16.7)	14 (12.1-15.8)	81.4 (70.6-92.2)	85.4 (75.2-99.0)
9 - <10 years	15.8 (13.5-18.2)	14 (11.4-16.5)	88 (74.5-101.4)	91.3 (78.9-110.1)
10 - <11 years	17.6 (13.3-22.0)	15.1 (10.6-20.4)	88.5 (71.1-106)	97.8 (77.6-123.9)

Sensitivity analysis

Sensitivity analyses using the DMI strategy instead of MID yielded no substantive differences in OCs for males and females (eFigures 1 and 2 and eTable 3). In complete-case analysis, OC estimates for males were also similar but the MID approach yielded slightly narrower CIs, particularly in infancy and late childhood (eFigures 3 and 4). In females, complete-case analysis generated similar OC estimates to the MID approach except at ages 8-years and older, where complete case analysis estimates were more variable with notably wider CIs, especially in the upper limit of OCs.

Interactive web-based tool

The Optimized Reference Assessment for Clinical Laboratory Evaluation for Ferritin (ORACLE-FER) Shiny app is an interactive web-tool designed to support the interpretation of ferritin concentrations for pediatric populations based on our study findings.[53] The app enables users to explore ferritin RIs and OIs, respectively generated from RCs and OCs, for children aged 2-weeks to 10-years. The user can enter the child's sex and birthdate or age to view age- and sex-specific RIs and OIs, visualize the optimality zone (area between lower/upper OCs), and compare values with WHO and ASH thresholds. The app supports multiple age formats (custom ranges, WHO age categories, standardized intervals) and displays an individual child's ferritin value relative to RCs and OCs.

DISCUSSION

In this study, we generated sex-specific ferritin RCs based on cross-sectional data from a healthy cohort of 4935 Canadian children (2322 females, 2613 males) aged 2-weeks to 10-years. We then applied pre-specified optimality criteria to define a sub-population at lower risk of ID to estimate sex-specific OCs. Given pronounced skewness of the distribution of ferritin measurements, we selected GAMLSS modeling, which accommodates asymmetry and enables

simultaneous modeling of the distribution's mean, variance, skewness, and kurtosis as smooth functions of age. Multiple imputation was used to handle missing data for optimality criteria. From RCs and OCs, we generated age- and sex-specific RIs and OIs with their corresponding CIs. We developed a web-based tool that enables users to obtain individualized RI and OI estimates for ferritin based on a child's exact age and sex and contextualizes the results with respect to WHO and ASH ID thresholds.

Our curves for ferritin show a dynamic relationship with age in pre-adolescents, with the lower-limit reaching a minimum of 5-6µg/L at 1.5 years and maximum of 15-17µg/L at 9-10 years. The similarity between ferritin RCs and OCs likely reflects the profile of the cohort, which was designed to represent a generally healthy population.[27] Furthermore, while our optimality criteria aimed to create a low risk sub-group, this is not equivalent to ideal iron status. Applying optimality criteria did not appear to substantially add incremental interpretive value.

Of four studies we identified that have developed pediatric ferritin RCs, two[9,46] used quantile regression and two used the Lambda-Mu-Sigma (LMS) method.[54,55] Three of the four studies[9,46,55] reported ferritin RCs below four years of age and no studies reported values below one year of age. In our study, the lowest ferritin values at 18-months of age were 5.9µg/L (90%CI, 5.3-6.5µg/L) from the OC for males and 6.8µg/L (90%CI, 6.1-7.5µg/L) from the OC for females. By comparison, the single study that provided 90%CI [9], reported 6.5µg/L (90%CI, 1.5-11.4µg/L) for males and 9.0µg/L (90%CI, 6.5-11.5µg/L) for females at 18-months. The other two studies that reported values at 18-months[46,55] reported values of 2.5 and 7.1µg/L for males and 9.0 and 11.5µg/L for females, respectively. Across studies, the lower limit is lower for males as compared to females, which may be due to more rapid growth in males.[56]

There is a developing consensus that the distribution-based lower RI limit (2.5th percentile) leads to underdiagnosis of ID.[26] The ASH threshold of ≤20µg/L is based on a single diagnostic accuracy study comparing ferritin measurements with bone marrow iron in 87 Malawian children (6-66 months) undergoing elective surgery.[26,57] More recent studies using a physiologically-based approach comparing ferritin with hemoglobin support this threshold.[12–17] Our findings suggest that applying a threshold of ≤20µg/L (rather than the lower RI limit or WHO threshold) will substantially increase the proportion of children diagnosed with ID. This is especially true for young children <3 years for whom prevalence may be higher than 50%.

The development of lower and upper limits of RIs and RCs has inherent limitations due to its distribution-based approach. Thresholds (also known as Clinical Decision Limits) are established based on clinically important outcomes. Using data from the same cohort, we have previously identified a ferritin threshold of 17-23.7µg/L for ID in children 1-3 years using a physiological outcome (hemoglobin) and a functional outcome (cognition).[15,53] However, ID thresholds do not provide an upper limit, which also has clinical utility. Furthermore, the ID threshold may vary across age, as seen in the continuous lower limit of the curves, suggesting a single pediatric threshold may not suffice.

Our interactive web-based tool is intended to complement, rather than replace existing current laboratory reporting standards which include reporting RIs. In addition, the tool contextualizes ID thresholds in terms of normative and optimal levels that vary across age and sex. However, the tool requires further validation by clinical and laboratory users. Future work will focus on expanding the underlying model using international datasets to improve generalizability.

Strengths and limitations

Strengths of our study include the large sample size of pre-adolescent children including those in the higher-risk age group of 1-3 years, use of data from a healthy community cohort, application of pre-specified optimality criteria, use of rigorous statistical methods following CLSI, use of multiple imputation to address missing data on optimality criteria, and contextualizing our findings alongside WHO and recently proposed ASH ID thresholds.

Our study has several limitations. Data were collected from a single urban Canadian center with exclusion of non-English-speaking families which may limit generalizability. Therefore, the curves and intervals should not be adopted internationally without local verification. It is notable that physiologically-based thresholds for young children from our Canadian cohort[14] are similar to those from U.S.[16-18] and 12 other countries.[19] Excluding children with a CRP>5 mg/L may have excluded some children with low-grade inflammation related to adiposity. In previous research in this cohort, we examined the complex relationship between ferritin, zBMI and CRP in young children and found no interaction between zBMI and CRP.[31] We lacked sufficient representation of neonates younger than 2-weeks and children older than 10 years. Iron disorders are rare in neonates, but ID is prevalent in adolescent females. Physiologically-based ID thresholds have been reported in a cohort of non-pregnant women including those as young as 15 years.[16,18] Missingness was assumed to be Missing at Random (MAR); however, the MAR assumption cannot be empirically confirmed from observed data alone, which remains an inherent limitation.

CONCLUSION

In this cross-sectional study, we derived ferritin RCs and OCs from a low-risk pre-adolescent population. In contrast to RIs, the curve approach addressed the limitations of age partitioning and sample size. However, the lower limit of curves and intervals were substantially

lower than physiologically-based thresholds to define ID. These findings further inform the current discussion regarding distributional versus physiological approaches to interpreting ferritin measurements and guiding clinical decision-making.

ACKNOWLEDGMENT SECTION

Author Contributions: Vid Bijelić had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: All authors.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Bijelić, Potter, Parkin, Hamid.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Bijelić, Hamid.

Obtained funding: Hamid, Parkin.

Administrative, technical, or material support: Potter, Hamid.

Supervision: Potter, Hamid.

Conflict of Interest Disclosures: Dr Parkin reported receiving grants from the Canadian Institutes of Health Research and the Hospital for Sick Children related to this work. The other authors have no conflicts to report.

Funding/Support: This work was supported by Natural Sciences and Engineering Research Council of Canada (grant number: RGPIN-2018-06693 to Dr. Hamid), and Canadian Institutes for Health Research (grant number: FRN 115059 to Dr. Parkin), and The Hospital for Sick Children Foundation (grant number: SP05-602 to Dr. Parkin).

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication

Non-author Contributions: We thank all participating children and families for their time and involvement in the TARGet Kids! primary care practice–based research network and all practice site physicians, research staff, collaborating investigators, trainees, methodologists, biostatisticians, data management personnel, laboratory management personnel, and advisory committee members.

Data Sharing Statement: The data that support the findings of this study are not publicly available as the participants did not give written consent for their data to be shared publicly.

Originality of Content: The authors verify that all information and materials in the manuscript are original.

References

1 World Health Organization. WHO guideline on use of ferritin concentrations to assess iron status in individuals and populations. Geneva: World Health Organization, 2020 https://iris.who.int/handle/10665/331505 (accessed Aug 5, 2024).

2 Garcia-Casal MN, Pasricha S-R, Martinez RX, Lopez-Perez L, Peña-Rosas JP. Serum or plasma ferritin concentration as an index of iron deficiency and overload. *Cochrane Database Syst Rev* 2021; **5**: CD011817.

3 Clinical and Laboratory Standards Institute (CLSI), editor. Defining, Establishing and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition. Wayne, Pa: Clinical and Laboratory Standards Institute, 2008.

4 Daly CH, Higgins V, Adeli K, Grey VL, Hamid JS. Reference interval estimation: Methodological comparison using extensive simulations and empirical data. *Clin Biochem* 2017; **50**: 1145–58.

5 Bijelić V, Potter B, Parkin PC, Momoli F, Liebman M, Hamid JS. Paediatric reference intervals and curves for haemoglobin and ferritin: protocol for a systematic review and meta-analysis. *BMJ Open* 2023; **13**: e073783.

6 Higgins V, Adeli K. Advances in pediatric reference intervals: from discrete to continuous. *J Lab Precis Med* 2018; **3**: 3–3.

7 Asgari S, Higgins V, McCudden C, Adeli K. Continuous reference intervals for 38 biochemical markers in healthy children and adolescents: Comparisons to traditionally partitioned reference intervals. *Clin Biochem* 2019; **73**: 82–9.

8 De Onis M. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007; **85**: 660–7.

9 Hall A, Bohn MK, Wilson S, Higgins V, Adeli K. Continuous reference intervals for 19 endocrine, fertility, and immunochemical markers in the CALIPER cohort of healthy children and adolescents. *Clin Biochem* 2021; **94**: 35–41.

10 De Onis M, Garza C, Habicht J-P. Time for a New Growth Reference. *Pediatrics* 1997; **100**: e8–e8.

11 de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. *Food Nutr Bull* 2004; **25**: S15-26.

12 Abdullah K, Birken CS, Maguire JL, *et al.* Re-Evaluation of Serum Ferritin Cut-Off Values for the Diagnosis of Iron Deficiency in Children Aged 12-36 Months. *J Pediatr* 2017; **188**: 287–90.

13 Sezgin G, Monagle P, Loh TP, *et al.* Clinical thresholds for diagnosing iron deficiency: comparison of functional assessment of serum ferritin to population based centiles. *Sci Rep* 2020; **10**: 18233.

14 Mei Z, Addo OY, Jefferds ME, Sharma AJ, Flores-Ayala RC, Brittenham GM. Physiologically based serum ferritin thresholds for iron deficiency in children and non-pregnant women: a US National Health and Nutrition Examination Surveys (NHANES) serial cross-sectional study. *Lancet Haematol* 2021; **8**: e572–82.

15 Mukhtarova N, Ha B, Diamond CA, Plumb AJ, Kling PJ. Serum Ferritin Threshold for Iron Deficiency Screening in One-Year-Old Children. *J Pediatr* 2022; **245**: 217–21.

16 Mei Z, Addo OY, Jefferds MED, *et al.* Comparison of Current World Health Organization Guidelines with Physiologically Based Serum Ferritin Thresholds for Iron Deficiency in Healthy Young Children and Nonpregnant Women Using Data from the Third National Health and Nutrition Examination Survey. *J Nutr* 2023; **153**: 771–80.

17 Addo OY, Mei Z, Jefferds MED, *et al.* Physiologically based serum ferritin thresholds for iron deficiency among women and children from Africa, Asia, Europe, and central America: a multinational comparative study. *Lancet Glob Health* 2025; **13**: e831–42.

18 Braat S, Pasricha S-R. Finding ferritin in the plateaus and valleys of iron deficiency. *Lancet Haematol* 2021; **8**: e539–40.

19 Parker ML, Storm S, Sholzberg M, Yip PM, Beriault DR. Revising Ferritin Lower Limits: It's Time to Raise the Bar on Iron Deficiency. *J Appl Lab Med* 2021; **6**: 765–73.

20 Naveed K, Goldberg N, Shore E, *et al.* Defining ferritin clinical decision limits to improve diagnosis and treatment of iron deficiency: A modified Delphi study. *Int J Lab Hematol* 2023; **45**: 377–86.

21 Truong J, Naveed K, Beriault D, Lightfoot D, Fralick M, Sholzberg M. The origin of ferritin reference intervals: a systematic review. *Lancet Haematol* 2024; **11**: e530–9.

22 Tang GH, Phillips Z, Poutanen S, *et al.* 'Raise the Bar', a Canadian intervention to combat disparities in the recognition and management of iron deficiency: Development of the intervention, feasibility and implementation. *Br J Haematol* 2025; : bjh.70150.

23 Troike KM, McShane AJ. Re-evaluating ferritin thresholds to diagnose iron deficiency. *Clin Biochem* 2025; **140**: 111020.

24 Wang D, Sra M, Glaeser-Khan S, *et al.* Cost-Effectiveness of Ferritin Screening Thresholds for Iron Deficiency in Reproductive-Age Women. *Am J Hematol* 2025; **100**: 1132–40.

25 Fan BE. Rethinking ferritin thresholds: towards a physiological-based definition of iron deficiency. *Lancet Glob Health* 2025; **13**: e1345–6.

26 The American Society of Hematology (ASH). ASH Draft Recommendations for Diagnosis of Iron Deficiency. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwj-h7em0LiQAxWf1fACHbWnMtcQFnoECBwQAQ&url=https%3A%2F%2Fwww.hematology.org%2F-%2Fmedia%2Fhematology%2Ffiles%2Feducation%2Fclinicians%2Fguidelines-quality%2Fida-dx-public-comment.pdf.

27 Carsley S, Borkhoff CM, Maguire JL, *et al.* Cohort Profile: The Applied Research Group for Kids (TARGet Kids!). *Int J Epidemiol* 2015; **44**: 776–88.

28 Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008; **61**: 344–9.

29 Oatley H, Borkhoff CM, Chen S, *et al.* Screening for Iron Deficiency in Early Childhood Using Serum Ferritin in the Primary Care Setting. *Pediatrics* 2018; **142**: e20182095.

30 Bayoumi I, Parkin PC, Birken CS, Maguire JL, Borkhoff CM, for the TARGet Kids! Collaboration. Association of Family Income and Risk of Food Insecurity With Iron Status in Young Children. *JAMA Netw Open* 2020; **3**: e208603.

31 Sypes EE, Parkin PC, Birken CS, *et al.* Higher Body Mass Index Is Associated with Iron Deficiency in Children 1 to 3 Years of Age. *J Pediatr* 2019; **207**: 198-204.e1.

32 Borkhoff SA, Parkin PC, Birken CS, Maguire JL, Macarthur C, Borkhoff CM. Examining the Double Burden of Underweight, Overweight/Obesity and Iron Deficiency among Young Children in a Canadian Primary Care Setting. *Nutrients* 2023; **15**: 3635.

33 Baker RD, Greer FR, The Committee on Nutrition. Diagnosis and Prevention of Iron Deficiency and Iron-Deficiency Anemia in Infants and Young Children (0–3 Years of Age). *Pediatrics* 2010; **126**: 1040–50.

34 Domellöf M, Braegger C, Campoy C, *et al.* Iron Requirements of Infants and Toddlers. *J Pediatr Gastroenterol Nutr* 2014; **58**: 119–29.

35 Unger SL, Fenton TR, Jetty R, Critch JN, O'connor DL. Iron requirements in the first 2 years of life. *Paediatr Child Health* 2019; **24**: 555–555.

36 Braat S, Fielding KL, Han J, *et al.* Haemoglobin thresholds to define anaemia from age 6 months to 65 years: estimates from international data sources. *Lancet Haematol* 2024; **11**: e253–64.

37 Statistics Canada. Low income cut-offs (LICOs) before and after tax by community size and family size, in current dollars. DOI:10.25318/1110024101-ENG.

38 Weltgesundheitsorganisation. Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age; methods and development. Geneva: WHO Press, 2006.

39 Namaste SM, Rohner F, Huang J, *et al.* Adjusting ferritin concentrations for inflammation: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. *Am J Clin Nutr* 2017; **106**: 359S-371S.

40 Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control* 1974; **19**: 716–23.

41 Stasinopoulos MD, Rigby RA, Heller GZ, Voudouris V, Bastiani FD. Flexible Regression and Smoothing: Using GAMLSS in R, 1st edn. Chapman and Hall/CRC, 2017 DOI:10.1201/b21973.

42 R Core Team. R: A Language and Environment for Statistical Computing. 2024. https://www.R-project.org/.

43 Eilers PHC, Marx BD. Flexible smoothing with B-splines and penalties. *Stat Sci* 1996; **11**. DOI:10.1214/ss/1038425655.

44 Dunn PK, Smyth GK. Randomized Quantile Residuals. *J Comput Graph Stat* 1996; **5**: 236–44.

45 Wozniak, M., Catomeris, P., Li, H., *et al.* Guidelines for the Use of Laboratory Tests for Iron Deficiency (CLP 002). Ontario Association of Medical Laboratories, 2024.

46 Hoq M, Matthews S, Karlaftis V, *et al.* Reference Values for 30 Common Biochemistry Analytes Across 5 Different Analyzers in Neonates and Children 30 Days to 18 Years of Age. *Clin Chem* 2019; **65**: 1317–26.

47 Parkin PC, Hamid J, Borkhoff CM, *et al.* Laboratory reference intervals in the assessment of iron status in young children. *BMJ Paediatr Open* 2017; **1**: e000074.

48 Buuren S van. Flexible imputation of missing data, Second edition. Boca Raton: Chapman & Hall/CRC, 2021.

49 Buuren SV, Groothuis-Oudshoorn K. **mice** : Multivariate Imputation by Chained Equations in *R*. *J Stat Softw* 2011; **45**. DOI:10.18637/jss.v045.i03.

50 Bijelić V, Bijelić M, Larock J, *et al.* Paediatric Reference Intervals and Curves for Haemoglobin Estimated Using Direct Methods: A Systematic Review and Meta-Analysis. *Int J Lab Hematol* 2025; : ijlh.14489.

51 Von Hippel PT. 4. Regression with Missing Ys: An Improved Strategy for Analyzing Multiply Imputed Data. *Sociol Methodol* 2007; **37**: 83–117.

52 Harrell Jr FE. rms: Regression Modeling Strategies. 2024. https://CRAN.R-project.org/package=rms.

53 Bijelić V. The Optimized Reference Assessment for Clinical Laboratory Evaluation for Ferritin (ORACLE-FER). 2025; published online Nov. https://svetrcocfer.shinyapps.io/ORACLE-FER/.

54 Rieger K, Vogel M, Engel C, *et al.* Reference intervals for iron-related blood parameters: results from a population-based cohort study (LIFE Child). *LaboratoriumsMedizin* 2016; **40**: 000010151520160019.

55 Wilson SM, Bohn MK, Madsen A, Hundhausen T, Adeli K. LMS-based continuous reference percentiles for 14 laboratory parameters in the CALIPER cohort of healthy children and adolescents. *Clin Chem Lab Med CCLM* 2023; **61**: 1105–15.

56 Davis SM, Kaar JL, Ringham BM, Hockett CW, Glueck DH, Dabelea D. Sex differences in infant body composition emerge in the first 5 months of life. *J Pediatr Endocrinol Metab* 2019; **32**: 1235–9.

57 Jonker FAM, Boele Van Hensbroek M, Leenstra T, *et al.* Conventional and novel peripheral blood iron markers compared against bone marrow in Malawian children. *J Clin Pathol* 2014; **67**: 717–23.

Supplemental tables and figures

eTable 1. Number of children used in estimations of reference and optimal ferritin curves by sex and standardized age groups.

Characteristic	Reference Curves		Optimal Curves	
	Female N = 2,322[1]	**Male N = 2,613**[1]	**Female N = 1,721**[1]	**Male N = 1,909**[1]
Age Group, n (%)				
2 weeks - 2 months	12 (0.5%)	22 (0.8%)	10 (0.6%)	13 (0.7%)
3 - 5 months	25 (1.1%)	25 (1.0%)	14 (0.8%)	13 (0.7%)
6 - 7 months	68 (2.9%)	65 (2.5%)	47 (2.7%)	41 (2.1%)
9 - 11 months	113 (4.9%)	145 (5.5%)	73 (4.2%)	80 (4.2%)
12 - 14 months	165 (7.1%)	185 (7.1%)	104 (6.0%)	107 (5.6%)
15 - 17 months	120 (5.2%)	140 (5.4%)	76 (4.4%)	98 (5.1%)
18 - 20 months	161 (6.9%)	165 (6.3%)	111 (6.4%)	106 (5.6%)
21 - 23 months	45 (1.9%)	62 (2.4%)	29 (1.7%)	36 (1.9%)
24 - 26 months	214 (9.2%)	208 (8.0%)	154 (8.9%)	163 (8.5%)
27 - 29 months	49 (2.1%)	53 (2.0%)	35 (2.0%)	41 (2.1%)
30 - 32 months	29 (1.2%)	43 (1.6%)	18 (1.0%)	21 (1.1%)
33 - 35 months	28 (1.2%)	44 (1.7%)	21 (1.2%)	30 (1.6%)
3 - <4 years	326 (14%)	319 (12%)	261 (15%)	252 (13%)
4 - <5 years	288 (12%)	354 (14%)	232 (13%)	288 (15%)
5 - <6 years	249 (11%)	266 (10%)	176 (10%)	203 (11%)
6 - <7 years	110 (4.7%)	146 (5.6%)	92 (5.3%)	117 (6.1%)
7 - <8 years	110 (4.7%)	102 (3.9%)	94 (5.5%)	83 (4.3%)
8 - <9 years	101 (4.3%)	112 (4.3%)	88 (5.1%)	90 (4.7%)
9 - <10 years	65 (2.8%)	89 (3.4%)	50 (2.9%)	69 (3.6%)
10 - <11 years	44 (1.9%)	68 (2.6%)	36 (2.1%)	58 (3.0%)

[1] n (%)

eTable 2. Age- and sex -specific reference intervals (RIs) and optimal intervals (OIs) with 90% confidence intervals (CIs) for children aged 0 - 131 months.

Month	Lower Limit		Upper Limit	
	Optimal Limit	Reference Limit	Optimal Limit	Reference Limit
Males				
0	31.3 (15.7-46.9)	30.6 (13.9-53.1)	620.8 (149.5-1092.1)	790.6 (466.6-1219.6)
1	25.9 (14.7-37.1)	23.2 (12.2-35.8)	484.3 (283.6-685.1)	570.7 (382.6-764.1)
2	21.5 (13.4-29.6)	17.5 (10.6-25)	383 (260.3-505.6)	414.8 (317.9-524.3)
3	18 (12.2-23.9)	13.6 (9.3-17.9)	306.4 (226.4-386.5)	309.3 (253.7-376.5)
4	15.3 (11-19.5)	10.8 (8-13.9)	248.3 (194-302.6)	237.5 (199.9-288)
5	13.1 (9.9-16.3)	8.9 (7-11.4)	204 (165.7-242.3)	190.3 (156.7-230.8)
6	11.4 (8.9-13.8)	7.7 (6.3-9.7)	170.2 (142.1-198.3)	155.8 (128.4-191.7)
7	10 (8.1-11.9)	6.9 (5.7-8.5)	144.4 (123-165.8)	133.3 (110.5-161.9)
8	9 (7.5-10.5)	6.4 (5.4-7.7)	124.4 (107.6-141.3)	118 (100.2-139)
9	8.2 (6.9-9.4)	6 (5.1-7.1)	108.9 (95.3-122.5)	106.7 (94.5-123.4)
10	7.5 (6.4-8.6)	5.8 (4.9-7)	96.7 (85.5-107.9)	98.3 (88.6-112.1)
11	7 (6.1-7.9)	5.7 (4.8-6.8)	87.1 (77.6-96.5)	92 (83.4-103.3)
12	6.6 (5.8-7.4)	5.5 (4.8-6.6)	79.4 (71.2-87.6)	86.3 (78-95.8)
13	6.3 (5.6-7.1)	5.4 (4.7-6.5)	73.4 (66.2-80.6)	81.3 (73.8-90.9)
14	6.1 (5.4-6.8)	5.3 (4.6-6.2)	68.6 (62.2-75.1)	76.4 (69.9-85.2)
15	6 (5.3-6.6)	5.2 (4.6-6)	64.9 (59-70.8)	72 (66.1-79.7)
16	5.9 (5.2-6.5)	5.1 (4.6-5.8)	62.1 (56.7-67.6)	68.2 (63-74.7)
17	5.8 (5.2-6.5)	5.1 (4.5-5.7)	60.1 (54.9-65.3)	65.2 (60.7-70.7)
18	5.9 (5.3-6.5)	5 (4.6-5.7)	58.7 (53.7-63.7)	62.6 (57.9-68.5)
19	6 (5.3-6.6)	5.1 (4.6-5.7)	58 (53.1-62.8)	61.1 (56.1-67)
20	6.1 (5.5-6.7)	5.2 (4.7-5.9)	57.7 (53-62.5)	60.5 (55.6-66.1)
21	6.3 (5.7-6.9)	5.4 (4.8-6)	57.9 (53.3-62.6)	60.4 (55.8-65.6)
22	6.5 (5.9-7.1)	5.6 (5.1-6.2)	58.4 (53.8-63.1)	61 (56.9-66.3)

23	6.7 (6.1-7.4)	5.8 (5.2-6.5)	59.2 (54.5-63.8)	61.8 (58-67.1)
24	7 (6.3-7.7)	6.1 (5.5-6.8)	60 (55.1-64.8)	62.6 (58.5-68)
25	7.2 (6.5-8)	6.3 (5.6-7.1)	60.7 (55.7-65.8)	63.4 (59-69.7)
26	7.5 (6.7-8.2)	6.5 (5.8-7.3)	61.4 (56.2-66.5)	63.9 (59.1-70.4)
27	7.7 (6.9-8.4)	6.7 (6-7.4)	61.7 (56.4-67)	64.1 (58.9-70.4)
28	7.9 (7.1-8.6)	6.9 (6.2-7.5)	61.9 (56.5-67.2)	64.1 (58.9-70.3)
29	8 (7.2-8.8)	7 (6.3-7.7)	61.9 (56.5-67.2)	63.6 (58.1-70.3)
30	8.2 (7.4-9)	7.2 (6.4-7.9)	61.8 (56.4-67.1)	63.1 (57.3-69.5)
31	8.3 (7.5-9.1)	7.3 (6.5-8.1)	61.6 (56.2-67)	62.9 (56.9-69.2)
32	8.4 (7.6-9.2)	7.4 (6.6-8.4)	61.4 (56.1-66.8)	62.7 (56.6-69.4)
33	8.5 (7.7-9.4)	7.5 (6.7-8.5)	61.3 (56.1-66.5)	62.5 (56.7-69.6)
34	8.7 (7.9-9.5)	7.6 (6.8-8.6)	61.3 (56.3-66.3)	62.4 (57.1-68.7)
35	8.8 (8-9.6)	7.8 (7.1-8.8)	61.4 (56.6-66.2)	62.4 (57.6-68.5)
36	9 (8.2-9.8)	8 (7.2-9)	61.6 (57-66.3)	62.9 (58.2-68)
37	9.2 (8.4-10)	8.2 (7.4-9.1)	62 (57.4-66.6)	63.3 (58.4-68)
38	9.4 (8.6-10.2)	8.4 (7.6-9.4)	62.4 (57.8-67)	63.8 (58.5-69.2)
39	9.6 (8.8-10.4)	8.7 (7.8-9.7)	63 (58.3-67.7)	64.5 (58.9-70.2)
40	9.8 (9-10.6)	8.9 (8-9.9)	63.7 (58.9-68.5)	65.2 (59.3-71.4)
41	10 (9.2-10.9)	9.2 (8.3-10.2)	64.5 (59.6-69.4)	66 (60.4-72.4)
42	10.3 (9.5-11.1)	9.4 (8.5-10.4)	65.4 (60.4-70.4)	66.7 (61.5-73.5)
43	10.5 (9.7-11.4)	9.6 (8.7-10.6)	66.3 (61.1-71.4)	67.5 (62.3-74.3)
44	10.8 (9.9-11.6)	9.8 (8.9-10.9)	67.1 (61.8-72.3)	68.1 (62.8-75.6)
45	11 (10.1-11.8)	10 (9-11.1)	67.8 (62.5-73.1)	68.4 (62.9-76)
46	11.2 (10.3-12)	10.1 (9.2-11.2)	68.4 (63.1-73.6)	68.6 (63.1-77)
47	11.3 (10.5-12.1)	10.2 (9.3-11.3)	68.7 (63.7-73.8)	68.4 (63.5-76.6)
48	11.4 (10.6-12.2)	10.3 (9.4-11.3)	69 (64.1-73.8)	68.3 (63.5-75.6)
49	11.6 (10.8-12.3)	10.4 (9.3-11.4)	69 (64.3-73.8)	68.1 (63.2-74.8)

50	11.6 (10.8-12.5)	10.4 (9.3-11.5)	69.1 (64.4-73.7)	68 (62.7-74.3)
51	11.7 (10.9-12.6)	10.5 (9.5-11.6)	69.1 (64.4-73.7)	68 (62.5-74.2)
52	11.8 (10.9-12.7)	10.6 (9.5-11.8)	69.1 (64.3-73.8)	67.8 (62.5-74)
53	11.9 (11-12.8)	10.7 (9.7-11.9)	69.1 (64.3-73.9)	67.9 (63-74.6)
54	12 (11.1-12.9)	10.8 (9.8-12)	69.2 (64.5-73.9)	68.3 (63.3-74.1)
55	12.1 (11.2-13)	11 (10-12.1)	69.3 (64.7-74)	68.9 (63.8-74.8)
56	12.2 (11.3-13.1)	11.2 (10.2-12.2)	69.5 (64.9-74.1)	69.6 (64-75.5)
57	12.3 (11.4-13.2)	11.3 (10.4-12.3)	69.6 (65-74.1)	70.4 (65.1-76.1)
58	12.3 (11.4-13.3)	11.5 (10.6-12.5)	69.6 (65.1-74.2)	71 (65.8-76.5)
59	12.4 (11.5-13.3)	11.6 (10.7-12.6)	69.6 (65.1-74.1)	71.4 (65.9-76.9)
60	12.5 (11.6-13.4)	11.7 (10.9-12.7)	69.5 (65.1-73.9)	71.5 (65.9-76.9)
61	12.5 (11.6-13.4)	11.8 (10.9-12.8)	69.3 (64.9-73.7)	71.5 (66.2-76.8)
62	12.5 (11.6-13.5)	11.8 (10.9-12.9)	69.1 (64.6-73.5)	71.4 (66.6-77)
63	12.5 (11.6-13.5)	11.9 (10.9-13)	68.8 (64.3-73.3)	71.3 (66.2-76.9)
64	12.6 (11.6-13.5)	11.9 (10.8-13.1)	68.5 (63.8-73.2)	71.2 (65.6-76.5)
65	12.6 (11.5-13.6)	11.9 (10.8-13.2)	68.2 (63.4-73.1)	71 (65.4-76.5)
66	12.6 (11.5-13.6)	12 (10.9-13.3)	68 (63-73)	71.1 (65.5-76.7)
67	12.6 (11.5-13.7)	12.1 (10.9-13.4)	67.8 (62.8-72.8)	71 (65.5-76.7)
68	12.7 (11.6-13.8)	12.1 (11-13.5)	67.7 (62.7-72.7)	71 (65.9-76.7)
69	12.7 (11.6-13.8)	12.2 (11.1-13.5)	67.6 (62.6-72.7)	71.3 (65.6-76.6)
70	12.8 (11.6-13.9)	12.3 (11.1-13.5)	67.7 (62.5-72.8)	71.4 (65.3-76.9)
71	12.8 (11.7-14)	12.3 (11.2-13.5)	67.7 (62.5-73)	71.4 (65.7-77.6)
72	12.9 (11.7-14.1)	12.3 (11.2-13.6)	67.9 (62.6-73.3)	71.3 (65.3-78.4)
73	13 (11.8-14.2)	12.3 (11.1-13.6)	68.1 (62.7-73.6)	70.8 (65.1-78)
74	13.1 (11.9-14.4)	12.3 (11.1-13.5)	68.4 (62.9-73.9)	70.4 (64.9-78.2)
75	13.3 (12-14.5)	12.4 (11-13.6)	68.8 (63.2-74.3)	69.9 (64.8-77.5)
76	13.4 (12.2-14.7)	12.3 (11-13.6)	69.2 (63.6-74.9)	69.4 (63.6-77.5)

77	13.6 (12.3-14.8)	12.3 (11-13.7)	69.7 (64-75.5)	69.3 (63.2-78.3)
78	13.7 (12.4-15)	12.4 (10.9-13.9)	70.3 (64.5-76.2)	69.3 (62.4-78.5)
79	13.9 (12.6-15.2)	12.5 (11-14.1)	71 (65.1-76.9)	69.3 (62.3-78.6)
80	14.1 (12.8-15.4)	12.7 (11.1-14.2)	71.8 (65.9-77.7)	70 (62.9-78.5)
81	14.3 (13-15.7)	12.9 (11.4-14.3)	72.6 (66.7-78.4)	70.9 (64.1-78.3)
82	14.5 (13.2-15.9)	13.1 (11.7-14.6)	73.4 (67.6-79.2)	72.1 (65.3-79)
83	14.7 (13.3-16.1)	13.4 (12-15)	74.2 (68.4-80)	73.5 (67.3-80.4)
84	14.9 (13.5-16.3)	13.7 (12.3-15.3)	75 (69.1-80.8)	74.7 (68.7-81.9)
85	15.1 (13.6-16.5)	14 (12.4-15.6)	75.7 (69.7-81.7)	75.8 (69.4-82.6)
86	15.2 (13.7-16.7)	14.2 (12.6-15.8)	76.3 (70.1-82.4)	77 (70.4-84.3)
87	15.3 (13.8-16.8)	14.3 (12.9-16)	76.7 (70.4-83)	77.4 (71-85.2)
88	15.4 (13.9-16.9)	14.5 (13-16.1)	77.1 (70.6-83.6)	78.2 (71.1-86.4)
89	15.4 (13.9-16.9)	14.6 (13.2-16.3)	77.3 (70.6-84)	78.5 (70.9-87.5)
90	15.4 (14-16.9)	14.7 (13.2-16.5)	77.4 (70.4-84.4)	78.6 (70.4-88.8)
91	15.5 (14-16.9)	14.8 (13.2-16.5)	77.5 (70.2-84.7)	78.6 (69.9-89.2)
92	15.5 (14-16.9)	14.8 (13.2-16.8)	77.4 (70-84.9)	78.7 (69.4-90.9)
93	15.4 (14-16.9)	14.8 (13.2-16.9)	77.4 (69.8-84.9)	78.7 (69.3-90.2)
94	15.4 (14-16.9)	14.8 (13.2-16.8)	77.3 (69.7-84.8)	78.6 (68.7-90.6)
95	15.4 (14-16.8)	14.8 (13.2-16.8)	77.2 (69.6-84.7)	78.5 (68.8-90.1)
96	15.4 (14-16.8)	14.8 (13.1-16.8)	77.1 (69.6-84.6)	78 (69.4-89.3)
97	15.4 (14-16.8)	14.8 (13-16.7)	77.1 (69.6-84.6)	77.7 (69.1-88.6)
98	15.4 (14-16.8)	14.8 (13.1-16.5)	77.1 (69.7-84.6)	77.5 (69.1-88.2)
99	15.4 (14-16.9)	14.9 (13.1-16.5)	77.3 (69.9-84.6)	77.4 (69.9-87.7)
100	15.5 (14-17)	15 (13.1-16.5)	77.5 (70.3-84.8)	77.5 (70.1-87.8)
101	15.6 (14.1-17.1)	15 (13.2-16.5)	77.9 (70.7-85.1)	77.7 (70.3-87.5)
102	15.7 (14.1-17.3)	15.2 (13.4-16.6)	78.4 (71.3-85.6)	78.2 (70.6-87.1)
103	15.8 (14.2-17.5)	15.3 (13.3-17)	79 (71.8-86.3)	78.6 (71.2-87.5)

104	16 (14.3-17.8)	15.5 (13.2-17.3)	79.8 (72.5-87.1)	79.2 (71-87.1)
105	16.2 (14.4-18)	15.7 (13.2-17.7)	80.6 (73.2-88.1)	80 (71.5-87.8)
106	16.4 (14.5-18.3)	15.8 (13.5-17.9)	81.6 (74-89.1)	81 (72-89.3)
107	16.7 (14.7-18.6)	16.1 (13.8-18.1)	82.6 (75-90.2)	82 (73.1-90.6)
108	16.9 (15-18.9)	16.4 (14.1-18.4)	83.7 (76-91.3)	82.8 (74.4-92)
109	17.2 (15.2-19.1)	16.6 (14.5-19)	84.7 (77-92.5)	84.3 (76-93)
110	17.4 (15.4-19.4)	16.8 (14.6-19.6)	85.8 (77.8-93.7)	85.6 (77.5-94.1)
111	17.7 (15.6-19.7)	17.1 (14.7-20.4)	86.7 (78.5-94.9)	87 (78.2-95.5)
112	17.9 (15.8-19.9)	17.4 (14.9-20.8)	87.5 (79-96)	87.9 (78.8-97.4)
113	18.1 (16-20.1)	17.6 (14.9-20.8)	88.1 (79.3-96.9)	88.4 (78.3-98.9)
114	18.2 (16.2-20.2)	17.7 (15.2-20.5)	88.4 (79.3-97.5)	88.7 (77.9-99.5)
115	18.3 (16.3-20.3)	17.8 (15.3-20.5)	88.5 (79.2-97.9)	88.3 (78.1-98.9)
116	18.3 (16.3-20.3)	17.8 (15.5-20.6)	88.4 (78.7-98)	88.2 (77.1-99)
117	18.3 (16.3-20.3)	17.8 (15.6-20.5)	87.9 (78-97.9)	87.9 (75.7-98.8)
118	18.2 (16.3-20.2)	17.7 (15.5-20.3)	87.2 (77.1-97.4)	87.3 (74.2-98)
119	18.1 (16.1-20)	17.5 (15.5-20)	86.2 (75.9-96.6)	86 (73.4-97.3)
120	17.9 (15.9-19.9)	17.4 (15.4-19.7)	85 (74.5-95.5)	84.9 (71.8-96.5)
121	17.7 (15.7-19.7)	17.2 (15.3-19.5)	83.5 (72.9-94.1)	83.6 (70.4-96.4)
122	17.4 (15.4-19.4)	17.1 (14.9-19.5)	81.9 (71.3-92.5)	82.4 (68.4-95)
123	17.1 (15.1-19.2)	16.9 (14.6-19.4)	80.2 (69.6-90.9)	81.3 (67.2-94.4)
124	16.9 (14.7-19)	16.7 (14.3-19.6)	78.6 (67.9-89.3)	80.2 (65.5-93.5)
125	16.6 (14.3-18.9)	16.5 (14.1-19.5)	76.9 (66.2-87.7)	78.8 (63.9-92.5)
126	16.4 (14-18.7)	16.3 (13.8-19.3)	75.4 (64.6-86.3)	77.7 (62.5-91.1)
127	16.1 (13.6-18.7)	16.1 (13.6-19.1)	74.1 (63.1-85)	76.8 (62-89.6)
128	16 (13.3-18.7)	16 (13.3-19.2)	72.8 (61.6-84.1)	75.5 (61.2-89.3)
129	15.8 (12.9-18.7)	16 (13.1-19.2)	71.7 (59.9-83.5)	74.6 (60.2-88.3)
130	15.7 (12.5-18.9)	16.1 (12.7-19.3)	70.7 (58.1-83.4)	74.1 (58.8-89.3)

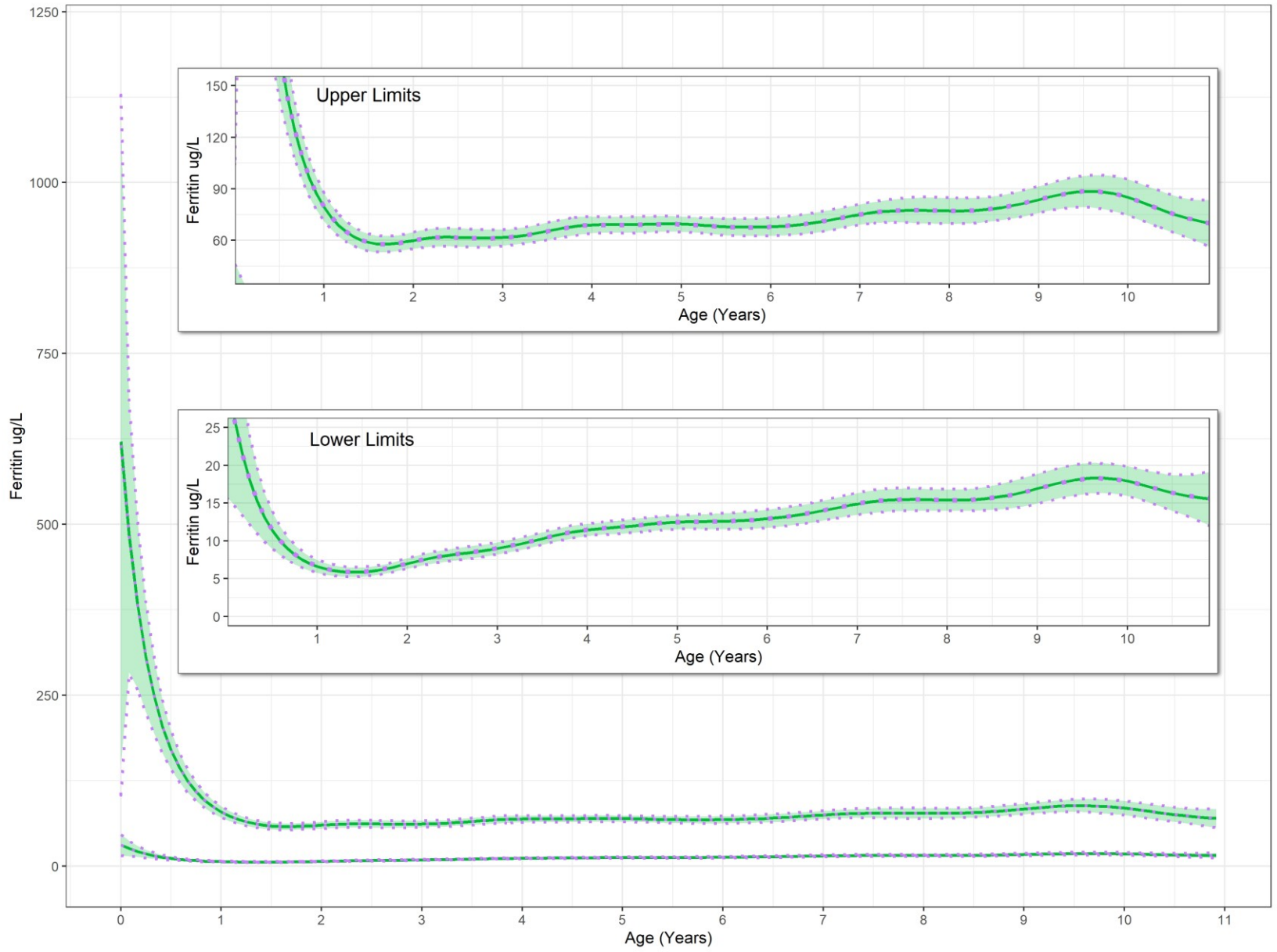
131	15.6 (12-19.2)	16.2 (12.2-19.9)	69.9 (56.1-83.6)	73.9 (55.9-90.9)
Females				
0	38 (6.5-69.4)	25.4 (14.1-54.5)	841 (424.6-1257.5)	733.9 (438.5-1251.7)
1	29.6 (8.7-50.4)	21.3 (12.9-39.3)	635.2 (380.5-889.9)	577.4 (379.7-892.1)
2	23.3 (9.5-37.2)	17.8 (11.7-29.1)	485 (329.7-640.4)	456.3 (328.1-641.2)
3	18.7 (9.4-28)	14.9 (10.4-22.1)	376 (280.1-472)	368.1 (278.2-475.6)
4	15.4 (9-21.7)	12.7 (9.4-17.4)	296.6 (235-358.3)	300.4 (238.4-372)
5	12.9 (8.5-17.2)	11 (8.4-13.9)	238.6 (195.9-281.3)	248.4 (202.9-292.9)
6	11 (7.9-14.2)	9.5 (7.6-11.7)	196.2 (163.9-228.5)	208.3 (174.4-243.6)
7	9.7 (7.4-12)	8.5 (7-10.1)	164.9 (138.8-190.9)	178.3 (152.1-205.6)
8	8.7 (6.9-10.4)	7.6 (6.4-9)	141.5 (119.7-163.2)	154.7 (132.9-174.7)
9	8 (6.6-9.4)	7 (5.9-8.1)	123.8 (105.4-142.2)	136.2 (118.1-152.8)
10	7.4 (6.3-8.6)	6.5 (5.6-7.5)	110.3 (94.6-126)	121.3 (106.8-137.2)
11	7 (6.1-8)	6.2 (5.3-7.1)	99.9 (86.4-113.4)	110.5 (97.9-124.7)
12	6.8 (5.9-7.7)	5.9 (5.1-6.8)	92 (80.2-103.7)	101.7 (91.1-114.5)
13	6.6 (5.8-7.4)	5.7 (4.9-6.5)	85.8 (75.5-96.1)	95.1 (85.4-105.6)
14	6.5 (5.8-7.3)	5.6 (4.9-6.4)	81.1 (72-90.3)	89.2 (81.2-98.5)
15	6.5 (5.8-7.3)	5.5 (4.8-6.4)	77.6 (69.3-85.8)	84.8 (77.8-93.4)
16	6.6 (5.9-7.3)	5.5 (4.9-6.3)	75 (67.3-82.6)	81.4 (74.8-89.5)
17	6.7 (6-7.4)	5.6 (4.9-6.3)	73.1 (65.9-80.3)	78.9 (71.8-86.2)
18	6.8 (6.1-7.5)	5.6 (5-6.3)	71.8 (64.9-78.7)	77.2 (70-83.8)
19	7 (6.3-7.7)	5.7 (5.1-6.4)	71 (64.2-77.8)	75.7 (68.6-81.9)
20	7.2 (6.5-8)	5.9 (5.2-6.5)	70.6 (63.9-77.2)	74.7 (67.9-80.6)
21	7.5 (6.7-8.2)	6.1 (5.3-6.7)	70.4 (63.8-76.9)	74 (67.5-79.9)
22	7.7 (7-8.5)	6.3 (5.5-7)	70.3 (63.8-76.9)	73.6 (67.2-79.5)
23	8 (7.2-8.8)	6.5 (5.6-7.2)	70.4 (63.8-77)	73.1 (66.9-79.6)
24	8.2 (7.4-9.1)	6.7 (5.8-7.5)	70.3 (63.7-77)	72.8 (66.5-79.4)
25	8.5 (7.6-9.3)	6.9 (5.9-7.7)	70.2 (63.4-77)	72.3 (65.9-79)

26	8.7 (7.7-9.6)	7 (6.1-7.9)	69.8 (62.9-76.6)	71.7 (65.1-78.4)
27	8.8 (7.9-9.7)	7.2 (6.3-8.1)	69.1 (62.2-75.9)	70.7 (64-77.9)
28	8.9 (7.9-9.8)	7.4 (6.5-8.3)	68.2 (61.4-75)	70 (62.8-77.2)
29	9 (8-9.9)	7.5 (6.6-8.5)	67.2 (60.5-74)	69.4 (61.8-76.5)
30	9 (8-10)	7.6 (6.7-8.6)	66.2 (59.5-72.9)	68.8 (61-75.7)
31	9 (8-10.1)	7.7 (6.9-8.7)	65.3 (58.7-71.9)	68.2 (60.3-75.1)
32	9.1 (8.1-10.1)	7.9 (7-8.9)	64.5 (58.1-70.9)	67.5 (59.3-74.4)
33	9.1 (8.1-10.2)	8 (7.1-9)	63.9 (57.8-70.1)	67.1 (59.1-73.7)
34	9.2 (8.2-10.2)	8.2 (7.3-9.1)	63.6 (57.7-69.4)	66.8 (59.4-73.2)
35	9.3 (8.4-10.3)	8.4 (7.5-9.2)	63.4 (57.9-69)	66.8 (60-73)
36	9.4 (8.5-10.4)	8.6 (7.7-9.5)	63.5 (58.1-68.8)	66.8 (60.7-72.8)
37	9.6 (8.6-10.5)	8.8 (7.9-9.8)	63.7 (58.4-69)	66.9 (61.1-73)
38	9.7 (8.8-10.7)	9 (8-10.1)	64 (58.6-69.5)	67.3 (61.3-73.3)
39	9.9 (8.9-10.9)	9.2 (8.1-10.3)	64.6 (59-70.1)	67.7 (61.6-73.6)
40	10.1 (9.1-11.1)	9.4 (8.3-10.5)	65.2 (59.5-70.9)	68.4 (61.9-73.9)
41	10.3 (9.3-11.3)	9.6 (8.5-10.6)	65.9 (60.1-71.7)	69.1 (62.4-74.7)
42	10.5 (9.5-11.5)	9.8 (8.7-10.8)	66.7 (60.8-72.6)	69.8 (63-75.6)
43	10.7 (9.7-11.7)	10 (8.8-11.1)	67.5 (61.6-73.5)	70.4 (63.7-76.5)
44	10.9 (9.9-12)	10.3 (9-11.3)	68.4 (62.3-74.4)	71 (64.5-77.6)
45	11.1 (10.1-12.2)	10.4 (9.2-11.4)	69.2 (63.2-75.2)	71.8 (65.5-78.5)
46	11.3 (10.3-12.3)	10.5 (9.4-11.6)	70 (64-75.9)	72.6 (66.1-79.5)
47	11.5 (10.5-12.5)	10.7 (9.6-11.8)	70.7 (64.9-76.5)	73.5 (66.6-80.1)
48	11.6 (10.6-12.6)	10.8 (9.8-11.9)	71.4 (65.7-77.1)	74.2 (67.7-80.6)
49	11.8 (10.8-12.8)	10.9 (9.8-12.1)	72 (66.3-77.7)	74.8 (68.4-81.7)
50	11.9 (10.9-12.9)	11 (9.9-12.2)	72.6 (66.8-78.3)	75.6 (68.9-82.2)
51	12 (10.9-13.1)	11.1 (9.9-12.2)	73.1 (67.2-78.9)	76.1 (69.7-82.9)
52	12.1 (11-13.2)	11.2 (10-12.3)	73.6 (67.6-79.5)	76.8 (70.4-83.6)

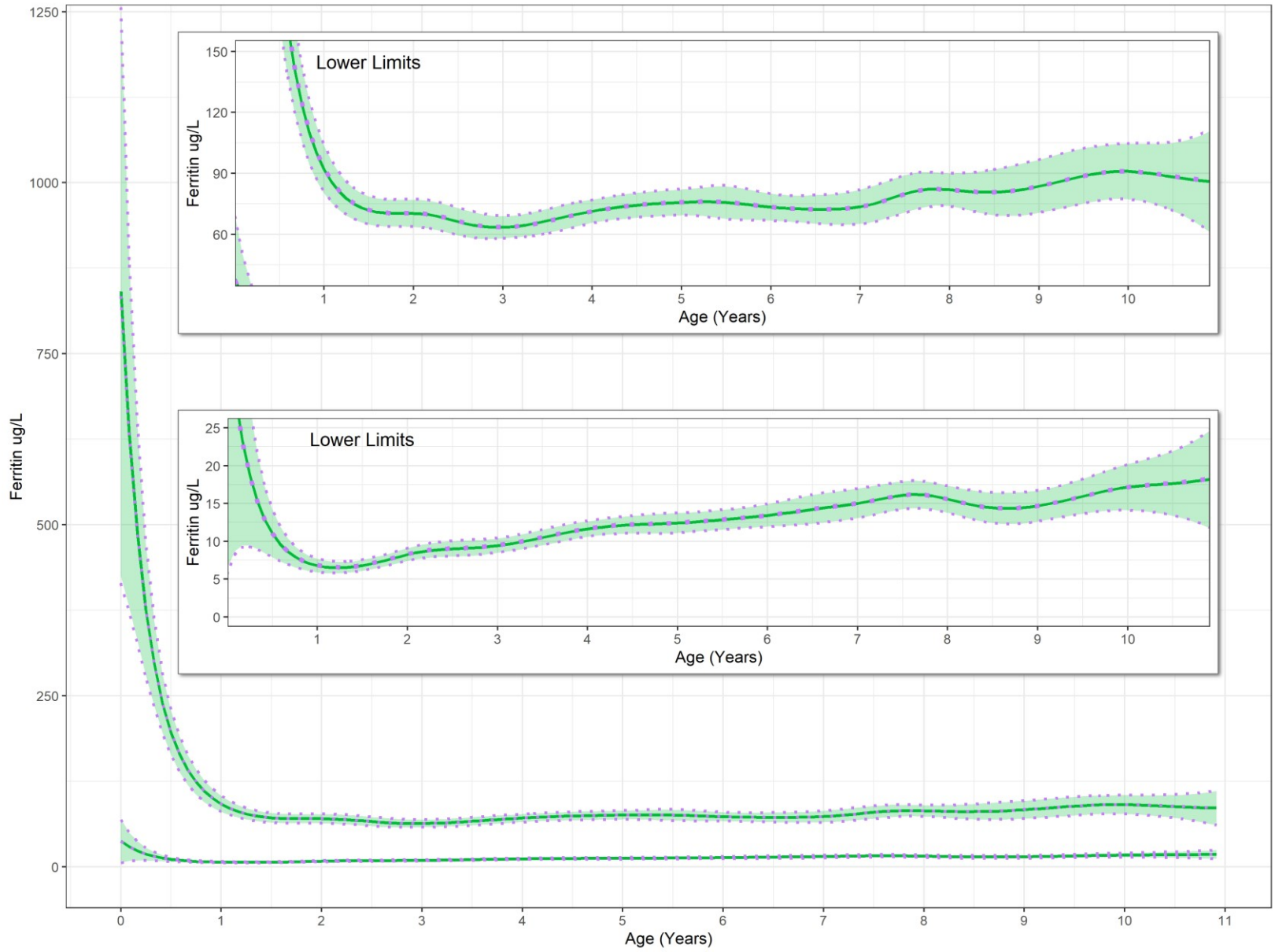
53	12.1 (11-13.3)	11.2 (10.1-12.3)	74 (67.9-80)	77.3 (70.6-84.2)
54	12.2 (11-13.4)	11.3 (10.1-12.5)	74.3 (68.2-80.4)	77.8 (71.1-85.1)
55	12.2 (11-13.4)	11.4 (10.1-12.6)	74.6 (68.5-80.7)	78.3 (71.5-85.5)
56	12.3 (11-13.5)	11.4 (10.1-12.6)	74.9 (68.8-81.1)	78.9 (72.2-86.1)
57	12.3 (11-13.6)	11.4 (10.1-12.6)	75.2 (68.9-81.4)	79.1 (73-86.9)
58	12.3 (11-13.6)	11.4 (10.1-12.7)	75.4 (69.1-81.6)	79.4 (73.8-87.9)
59	12.4 (11-13.7)	11.4 (10-12.7)	75.6 (69.3-81.9)	79.9 (73.8-88.6)
60	12.4 (11.1-13.7)	11.5 (10.1-12.7)	75.7 (69.4-82.1)	80.1 (73.7-89.7)
61	12.5 (11.2-13.8)	11.5 (10.2-12.7)	75.9 (69.4-82.3)	80.2 (73.6-90.1)
62	12.5 (11.3-13.8)	11.6 (10.3-12.8)	76 (69.3-82.7)	80.7 (73.3-90.3)
63	12.6 (11.3-13.9)	11.6 (10.3-13)	76 (69-83.1)	80.6 (73-90.3)
64	12.7 (11.4-14)	11.7 (10.3-13.1)	76 (68.6-83.5)	80.7 (72.7-90.7)
65	12.8 (11.5-14.1)	11.8 (10.4-13.2)	75.9 (68.1-83.7)	80.7 (72.1-91.1)
66	12.9 (11.5-14.2)	11.9 (10.5-13.2)	75.7 (67.7-83.6)	80.5 (72.3-91)
67	12.9 (11.6-14.3)	11.9 (10.6-13.2)	75.4 (67.5-83.2)	80.5 (72.2-90.6)
68	13 (11.7-14.3)	12 (10.6-13.3)	75 (67.3-82.7)	80.1 (72.5-89.8)
69	13.1 (11.8-14.4)	12.1 (10.6-13.3)	74.6 (67.2-82)	79.7 (72.6-89.2)
70	13.2 (11.8-14.6)	12.1 (10.6-13.5)	74.2 (67.1-81.3)	79 (72.1-88.7)
71	13.3 (11.9-14.7)	12.2 (10.7-13.6)	73.8 (66.9-80.6)	78.3 (71.5-87.9)
72	13.4 (11.9-14.9)	12.2 (10.7-13.7)	73.4 (66.8-80)	77.5 (71.3-86.8)
73	13.5 (12-15.1)	12.3 (10.8-13.8)	73.1 (66.6-79.6)	76.6 (69.9-85.6)
74	13.7 (12-15.3)	12.3 (10.8-14)	72.8 (66.4-79.3)	75.7 (69-84.1)
75	13.8 (12.1-15.5)	12.3 (10.8-14.2)	72.6 (66.1-79.1)	75 (68.2-82.5)
76	14 (12.2-15.7)	12.3 (10.8-14.5)	72.5 (65.8-79.1)	74.1 (66.2-81.2)
77	14.1 (12.3-15.9)	12.3 (10.9-14.7)	72.4 (65.5-79.2)	73 (64.6-80.4)
78	14.2 (12.3-16.1)	12.3 (11-14.8)	72.3 (65.2-79.3)	72.4 (63.4-80.4)
79	14.4 (12.4-16.3)	12.4 (10.9-15)	72.3 (65-79.5)	72 (62.7-80.4)

80	14.5 (12.5-16.4)	12.4 (11-15.2)	72.3 (64.8-79.8)	71.8 (62.2-80.9)
81	14.6 (12.7-16.6)	12.6 (10.9-15.3)	72.4 (64.7-80.1)	71.8 (62.2-80.5)
82	14.7 (12.8-16.7)	12.7 (11.1-15.6)	72.6 (64.7-80.6)	72.4 (62.2-81.1)
83	14.9 (12.9-16.8)	12.9 (11.1-15.7)	73 (64.9-81.2)	72.9 (62.6-82.1)
84	15 (13.1-17)	13.1 (11.2-15.8)	73.6 (65.1-82)	73.9 (62.9-82.9)
85	15.2 (13.2-17.1)	13.4 (11.4-16.3)	74.3 (65.6-83)	74.7 (64.2-84.2)
86	15.4 (13.4-17.3)	13.7 (11.9-16.5)	75.2 (66.4-84.1)	75.9 (65.7-85.3)
87	15.6 (13.7-17.5)	14.1 (12.1-16.8)	76.4 (67.4-85.3)	77.4 (67.3-87.1)
88	15.8 (13.9-17.6)	14.4 (12.3-17.1)	77.6 (68.5-86.6)	79.3 (69.5-88.9)
89	15.9 (14.1-17.8)	14.7 (12.7-17.5)	78.8 (69.7-87.9)	81.6 (71.3-90.9)
90	16.1 (14.2-17.9)	14.9 (13.1-17.8)	79.9 (70.8-89.1)	82.9 (73.1-92.9)
91	16.2 (14.3-18)	15.2 (13.4-18)	81 (71.9-90)	84.6 (75.2-94.7)
92	16.2 (14.4-18)	15.3 (13.5-18)	81.7 (72.8-90.6)	85.9 (76.9-95.9)
93	16.1 (14.3-17.9)	15.4 (13.5-17.8)	82.2 (73.6-90.7)	86.4 (77.9-97.1)
94	16 (14.2-17.8)	15.3 (13.4-17.6)	82.3 (74.1-90.5)	86.7 (78.6-97.3)
95	15.8 (14-17.6)	15.2 (13.3-17.3)	82.2 (74.2-90.3)	86.5 (79.1-96.4)
96	15.6 (13.8-17.4)	15 (13.1-17.1)	82 (73.9-90)	86.3 (78.6-95.9)
97	15.3 (13.5-17.1)	14.7 (13-16.7)	81.6 (73.3-90)	86 (77.2-95.9)
98	15.1 (13.3-16.9)	14.4 (12.7-16.2)	81.3 (72.4-90.1)	85.5 (77.1-96.3)
99	14.9 (13-16.8)	14.2 (12.3-15.9)	81 (71.5-90.4)	85.3 (76.9-97.2)
100	14.7 (12.8-16.6)	13.9 (12.1-15.8)	80.8 (70.7-90.8)	84.9 (76.1-98.1)
101	14.5 (12.6-16.5)	13.8 (12-15.7)	80.7 (70-91.4)	84.5 (75.3-99.3)
102	14.4 (12.4-16.5)	13.7 (11.9-15.6)	80.7 (69.5-92)	84.3 (74.5-99.4)
103	14.4 (12.3-16.4)	13.6 (11.7-15.4)	80.9 (69.1-92.7)	84.6 (73.8-101.2)
104	14.3 (12.3-16.4)	13.6 (11.7-15.4)	81.2 (69-93.4)	85.1 (73.7-101.1)
105	14.4 (12.3-16.4)	13.5 (11.7-15.5)	81.6 (69.1-94.2)	85.5 (73.3-100.5)
106	14.4 (12.4-16.5)	13.5 (11.5-15.5)	82.2 (69.4-95)	86.3 (72.9-100.6)

107	14.5 (12.5-16.6)	13.6 (11.6-15.5)	82.9 (69.9-95.9)	86.8 (73.2-103.1)
108	14.7 (12.6-16.7)	13.6 (11.6-15.5)	83.7 (70.6-96.7)	87.2 (74.3-104.8)
109	14.8 (12.8-16.9)	13.6 (11.6-15.6)	84.5 (71.3-97.7)	88 (74.8-105.6)
110	15 (13-17.1)	13.7 (11.6-15.6)	85.4 (72-98.7)	88.7 (75.5-107.9)
111	15.3 (13.2-17.4)	13.8 (11.5-15.7)	86.2 (72.7-99.8)	89.7 (76.9-109.7)
112	15.5 (13.3-17.6)	13.9 (11.5-15.8)	87.1 (73.5-100.7)	90.2 (78.2-109.8)
113	15.7 (13.5-17.9)	14 (11.5-15.9)	88 (74.3-101.6)	91.1 (78.8-108.7)
114	16 (13.6-18.3)	14 (11.4-16.2)	88.8 (75.1-102.4)	91.6 (79.7-108.8)
115	16.2 (13.8-18.6)	14.1 (11.3-16.6)	89.5 (75.9-103)	92.2 (80.1-110.3)
116	16.4 (13.9-18.9)	14.3 (11.2-17)	90.1 (76.6-103.6)	93.3 (80.8-111.8)
117	16.6 (14-19.3)	14.3 (11.1-17.5)	90.6 (77.1-104)	93.9 (81.7-112.8)
118	16.8 (14.1-19.6)	14.5 (11.3-18.3)	90.9 (77.4-104.4)	94.3 (83.1-114.8)
119	17 (14.1-19.9)	14.6 (11.3-18.8)	91 (77.5-104.6)	95.2 (83.1-116.8)
120	17.1 (14.1-20.2)	14.7 (11.3-19.2)	91 (77.3-104.6)	96.1 (82.7-117)
121	17.3 (14.1-20.5)	14.8 (11.5-19.4)	90.8 (76.9-104.6)	96.7 (82.1-117.3)
122	17.4 (14-20.7)	14.9 (11.3-19.2)	90.4 (76.3-104.5)	97.1 (82-118.7)
123	17.4 (13.9-21)	15 (11.1-19)	89.9 (75.5-104.4)	97.9 (81.1-119.8)
124	17.5 (13.8-21.3)	15 (10.9-19.4)	89.4 (74.4-104.5)	98.2 (80.8-121)
125	17.6 (13.6-21.6)	14.9 (10.6-19.9)	88.8 (73-104.6)	98.5 (81-122)
126	17.6 (13.4-21.9)	15.1 (10.5-20.2)	88.3 (71.5-105)	98.9 (78.7-122.7)
127	17.7 (13.1-22.3)	15.2 (10.3-20.5)	87.7 (69.9-105.6)	98.3 (76.8-125.7)
128	17.8 (12.9-22.8)	15.3 (10.1-21.2)	87.2 (68-106.5)	97.8 (74.8-127.3)
129	17.9 (12.5-23.3)	15.3 (9.9-21.4)	86.7 (65.8-107.6)	97.8 (72.8-129.1)
130	18.1 (12.2-24)	15.3 (9.7-22)	86.3 (63.5-109.1)	98.1 (70.6-131.7)
131	18.2 (11.7-24.7)	15.3 (9.6-23.3)	85.9 (61-110.9)	98 (67.7-134)



eFigure 1. Ferritin optimal curves for males aged between 2 weeks and 10 years, estimated using two different imputation strategies: multiple imputation first then delete (MID) in green and delete then multiple imputation (DMI) in purple. The 90% CIs for the lower and upper limits of the optimal curves are presented using dotted lines. The two row inserts (row panels) show lower and upper limits separately, for better visualization.

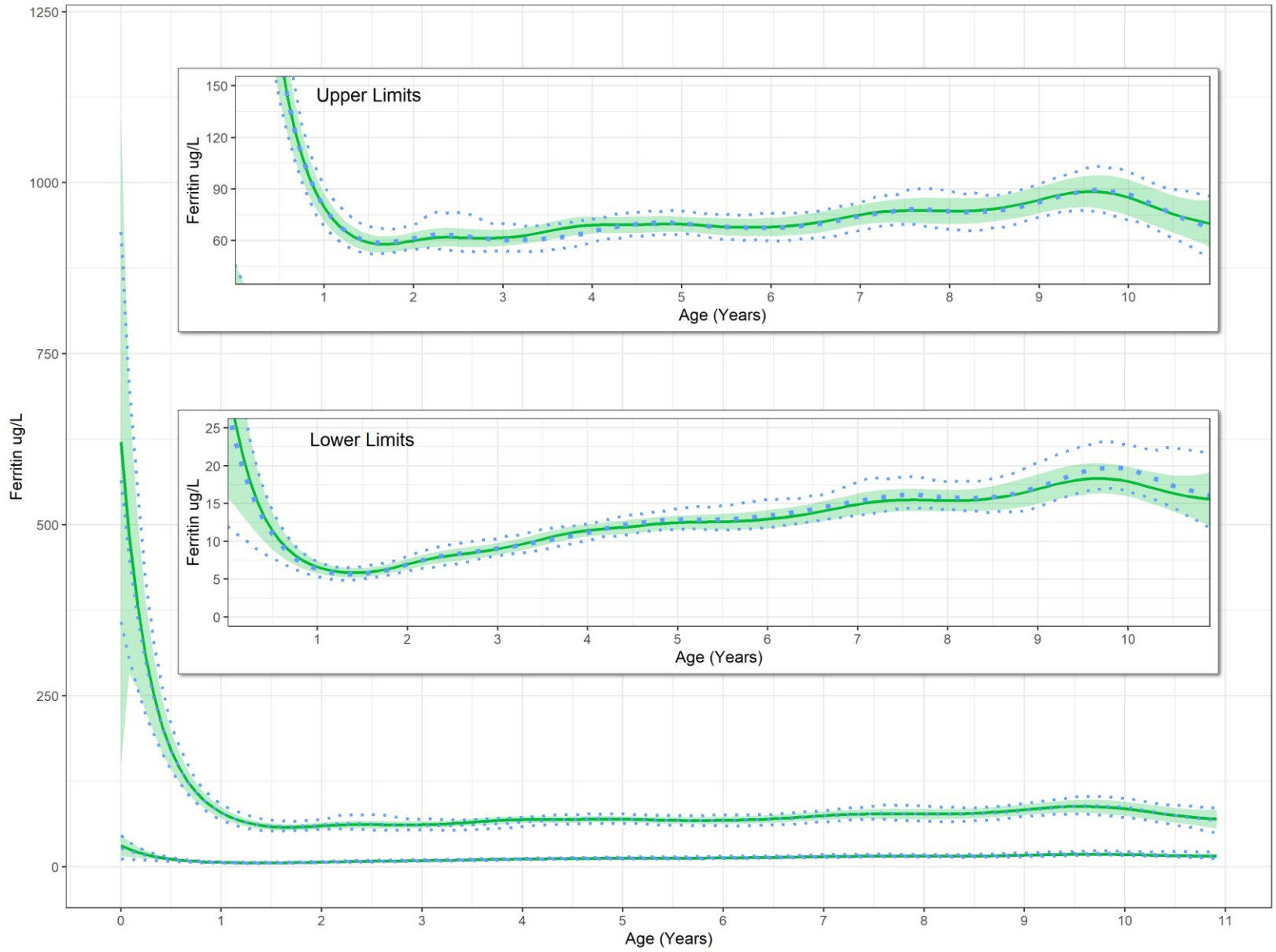


eFigure 2. Ferritin optimal curves for females aged between 2 weeks and 10 years, estimated using two different imputation strategies: multiple imputation first then delete (MID) in green and delete first then multiple imputation (DMI) in purple with 90% confidence intervals presented using dotted lines. The wo inserts shown in the two row panels provide the lower and upper limits of the optimal curves separately.

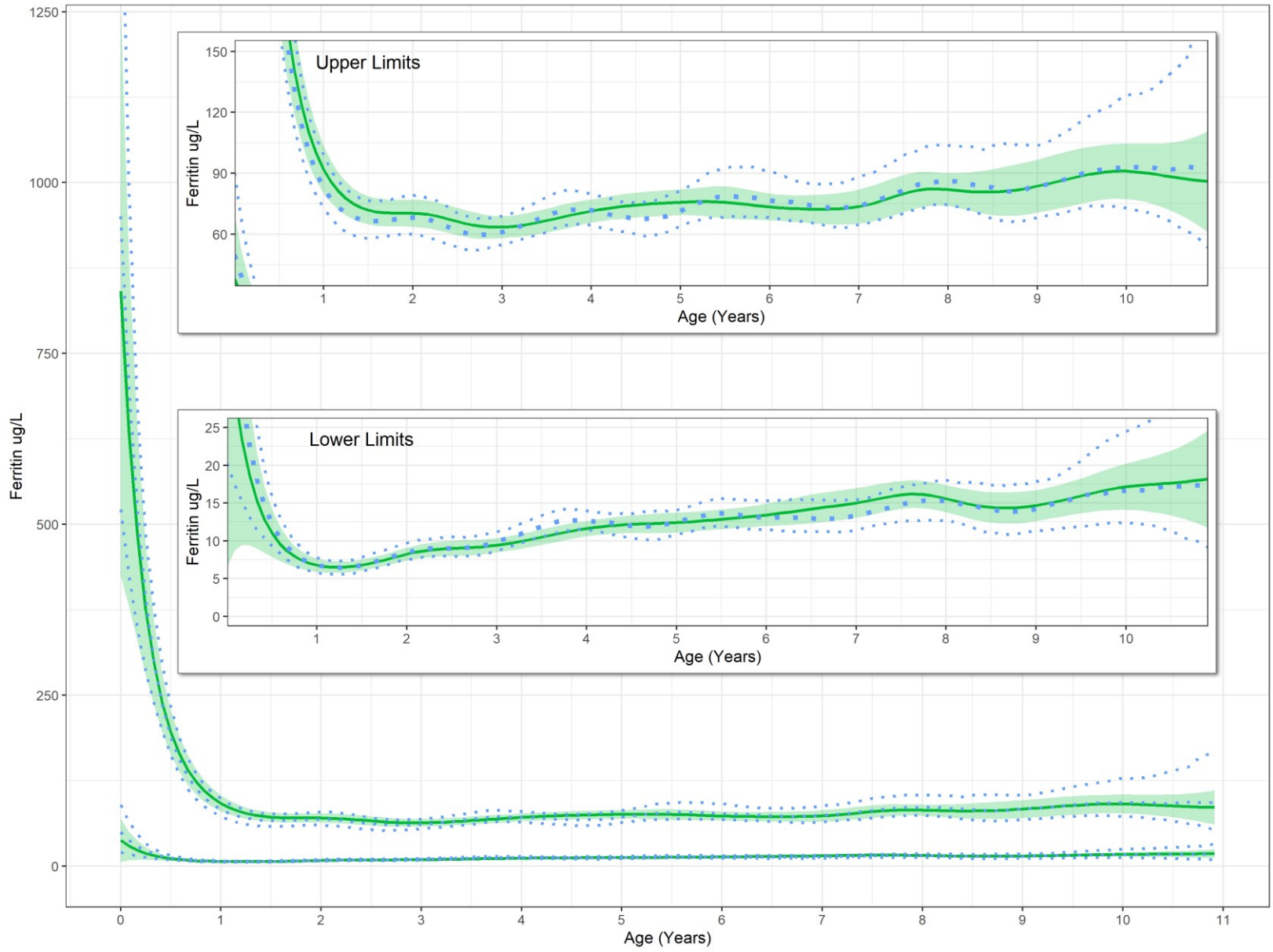
eTable 3. Characteristics of study participants in imputed and complete case datasets used for optimal curve estimation, prior to excluding those not meeting optimality criteria

Characteristics	**All participants with ferritin values[3] N = 4,935**	**Participants in primary MID analysis[4] N = 4,935**	**Participants in DMI sensitivity analysis[5] N = 4,935**	**Participants in complete case sensitivity analysis[6] N = 3,414**
Age in months, median (IQR)	37 (18, 62)	37 (18, 62)	37 (18, 62)	36 (18, 61)
Sex, n (%)				
Female	2,322 (47.1)	2,322 (47.1)	2,322 (47.1)	1,593 (46.7)
Male	2,613 (52.9)	2,613 (52.9)	2,613 (52.9)	1,821 (53.3)
Ferritin (ug/L), median (IQR)	29 (20, 41)	29 (20, 41)	29 (20, 41)	29 (20, 41)
Missing, n (%)	0 (0)	0 (0)	0 (0)	0 (0)
CRP (mg/L), median (IQR)	0.3 (0.2, 0.6)	0.3 (0.2, 0.6)	0.3 (0.2, 0.6)	0.2 (0.2, 0.5)
Missing, n (%)	0 (0)	0 (0)	0 (0)	0 (0)
Low income cut-off (LICO), n (%)				
Above LICO	3,985 (90.5)	4,464 (90.5)	4,461 (90.4)	3,079 (90.2)
Below LICO	417 (9.5)	471 (9.5)	474 (9.6)	335 (9.8)
Missing	533 (10.8)	0 (0)	0 (0)	0 (0)
Prematurity[1] n (%)				
Not premature	4,555 (96.0)	4,734 (95.9)	4,719 (95.6)	3,248 (95.1)
Premature	189 (4.0)	201 (4.1)	216 (4.4)	166 (4.9)
Missing	191 (3.9)	0 (0)	0 (0)	0 (0)
Low birthweight[1] n (%)				
Not low birthweight	4,678 (96.3)	4,746 (96.2)	4,744 (96.1)	3,263 (95.6)
Low birthweight	181 (3.7)	189 (3.8)	191 (3.9)	151 (4.4)
Missing	76 (1.5)	0 (0)	0 (0)	0 (0)
Obese/Underweight[2] n (%)				
Not Obese/Underweight	4,447 (92.0)	4,539 (92.0)	4,540 (92.0)	3,158 (92.5)
Obese/Underweight	385 (8.0)	396 (8.0)	395 (8.0)	256 (7.5)
Missing	103 (2.1)	0 (0)	0 (0)	0 (0)
Anemia, n (%)				
No anemia	4,189 (92.5)	4,569 (92.6)	4,556 (92.3)	3,165 (92.7)
Anemia	341 (7.5)	366 (7.4)	379 (7.7)	249 (7.3)
Missing	405 (8.2)	0 (0)	0 (0)	0 (0)

[1]Applicable when ferritin measured in first 24 months of life
*[2]Above or Below zBMI +/- 1.96*SD, where zBMI is BMI z-score relative to WHO reference population adjusted for age and sex*
[3]Sample for reference curve estimation, also included in Table 1 of main paper
[4]Missing data for optimality criteria variables imputed with multiple imputation then deletion (MID) approach (see Methods), also included in Table 1 of main paper for primary analysis to estimate optimality curves
[5]Missing data for optimality criteria variables imputed with deletion then multiple imputation (DMI) approach (see Methods), for sensitivity analysis to estimate optimality curves
[6]Participants with missing data on optimality criteria excluded, for sensitivity analysis to estimate optimality curves



eFigure 3. Ferritin optimal curves for males aged between 2 weeks and 10 years, comparing the estimates obtained using the multiple imputation then delete (MID) approach presented in green and complete case analysis presented in sky blue. The 90% CIs are presented using dotted lines. Two inserts show lower (lower inset) and upper (upper inset) optimal curve limits magnified.



eFigure 4. Ferritin optimal curves for females aged between 2 weeks and 10 years, comparing estimated derived using the multiple imputation then delete (MID) approach (green) and complete case analysis (sky blue), with dotted lines representing the 90% confidence intervals. The two inserts in the top row panels provide the lower and upper limits, separately.

CHAPTER 7: INTEGRATED DISCUSSION

Over seventy percent of medical decisions are informed by laboratory test results [1], making RIs an essential clinical tool used by both physicians and laboratory professionals when interpreting these values.[1,2] The International Federation of Clinical Chemistry and Laboratory Medicine has published a comprehensive series of recommendations for the establishment of RIs, including detailed protocols for selecting reference populations specifying exclusion criteria, and methods for partitioning the reference population by relevant factors such as age and sex. These steps, considered to be gold standard procedures for establishing RIs, are also outlined by the CLSI guidelines.[2] While these guidelines generally recommend that each laboratory establish its own RIs to reflect local population characteristics, full adherence to these recommendations is often resource demanding.[3]

In pediatrics, the challenge is even greater. Children undergo rapid physiological changes from birth through adolescence, requiring more age partitions than adults in RI estimation to accurately reflect biomarker variation.[2,4,5] This results in sparse data especially in infancy and early childhood, making it difficult to meet the sample size requirements recommended by CLSI guidelines.[2] Moreover, acquiring blood samples from pediatric populations has consistently proven to be challenging.[6] RCs, which model age as a continuous variable, offer a potential solution by enabling RI estimation at any age while eliminating the need for partitions with respect to age. RCs are, therefore, better suited to capturing the dynamic, non-linear trajectories of physiological biomarkers across development.[7–9] Despite this, a national survey of Canadian clinical laboratories revealed significant variability and lack of harmonization in pediatric RIs across institutions and a lack of use of RCs.[10,11]

In this thesis we aimed to characterize the current state of hemoglobin and ferritin RIs and RCs for the pediatric population, address key methodological and computational gaps, and offer comparative evaluations to inform their clinical interpretation and use. Hematological tests for iron deficiency are among the most commonly ordered tests in pediatric care, often obtained during routine visits.[12,13] Their clinical importance is highlighted by the fact that iron deficiency affects more than 750 million children worldwide.[13–15] Ferritin is often ordered together with hemoglobin for assessing iron status in children.[13,15,16] Hence, we focused on hemoglobin and ferritin, key biomarkers of iron status, to better understand the broader picture of pediatric RIs in Canada and globally. We began by synthesizing existing evidence on hemoglobin and ferritin RIs and RCs, then developed new RCs for both biomarkers from which sex-specific RIs can be generated at any age or age interval, and finally we identified key criteria for optimal biomarker status and utilized them to establish OCs and OIs. Alongside this, we established methodological frameworks to support the work and developed web-based computational and graphical tools that allow further investigation of global variations, enable comparative analysis and facilitate clinical interpretations.

7.1 Summary of manuscript 2 (Chapter 3): Pediatric reference intervals and curves for hemoglobin estimated using direct methods: A systematic review and meta-analysis

This manuscript presents a systematic review and meta-analysis of pediatric RIs and RCs for hemoglobin. Hemoglobin measurement is a fundamental laboratory test for assessing individual and population-level health.[17–19] To interpret hemoglobin test results, laboratories provide RIs with lower (2.5th percentile) and upper (97.5th percentile) limits, typically stratified by age and sex.[2] RCs that treat age as a continuous variable offer a more precise understanding of hemoglobin levels across pediatric developmental stages relative to RIs.[20] There is substantial

variability in the estimation of hemoglobin RIs across regions and countries, with notable discrepancies between RIs used in pediatric laboratories in Canada and globally.[10,11,17]

In this paper, we presented the results of a systematic review and meta-analysis that synthesized existing pediatric RIs and RCs for hemoglobin. The paper also investigated heterogeneity and identified its potential sources.[21] Out of 9 123 studies screened, 48 met the inclusion criteria, encompassing data from 63 529 male and 59 969 female participants across 25 countries and four continents, published between 1938 and 2023. The analysis revealed inconsistencies in age partitioning and age interval lengths among studies. Thirteen studies reporting RIs and two studies reporting RCs were included in the meta-analysis. Due to high heterogeneity, pooled estimates for the 0-3 months age group could not be generated. For children aged 3 months and older, both lower and upper RI limits generally increased with age, ranging approximately from 100 to 130 g/L and from 130 to 150 g/L, respectively. Notably, the lower RI limits in many studies differed substantially from WHO anemia thresholds. A lower RI limit compared to WHO thresholds was expected, as WHO anemia thresholds are based on the 5th percentile while RIs use the 2.5th percentile. However, instances where RI lower limits were higher than WHO thresholds were less expected.[21]

We concluded that there is substantial heterogeneity in pediatric hemoglobin RIs and RCs, which hindered quantitative synthesis for some of the age groups, underscoring the need for more rigorously and consistently developed estimates that can be used globally alongside WHO thresholds to define anemia. The paper advocated for future research focusing on establishing RIs for the youngest children and the exploration of percentile curves to provide continuous hemoglobin charts.

7.2 Summary of manuscript 3 (Chapter 4): Pediatric reference intervals and curves for ferritin estimated using direct methods: A systematic review and meta-analysis

This manuscript presents a systematic review and meta-analysis of pediatric RIs and RCs for ferritin. Ferritin is the most widely used biomarker for assessing iron status in children.[13,15,16] Concentration levels of ferritin represent a non-invasive alternative to bone marrow and kidney biopsy for identifying ID and iron overload respectively.[16] In clinical practice, ferritin levels are interpreted using RIs derived from healthy populations or in reference to clinical decision limits (CDLs), similar to WHO ferritin thresholds[16], derived from expert consensus.[2,15] The systematic review presented in this chapter synthesized pediatric RIs and RCs for ferritin, examined methodological quality of studies establishing RIs and RCs, quantified heterogeneity across studies, and identified sources of heterogenity.[22]

The systematic review included 20 eligible studies published between 1976 and 2023, representing data from 26,006 healthy children (13,143 male and 12,863 female) across North America, Europe, and Asia. RIs from 8 studies and a RC from one study that adhered to the CLSI guidelines were included in the meta-analysis. The pooled lower RI limits for ferritin showed a U-shaped trajectory, declining from infancy to toddlerhood and increasing thereafter, with a low point at 6-9 μg/L between 24-33 months of age. The u-shaped trajectory is consistent with results of the relationship of ferritin with age for a healthy pediatric population estimated in our serum ferritin RCs and OCs study presented in Chapter 6. Substantial heterogeneity was identified in both upper and lower RI limits, especially in younger children and for females. Factors contributing to heterogeneity included inconsistent age partitioning, geographic location, and laboratory analyzer types. Included studies failed to assess or exclude participants with inflammation, anemia, or known risk factors for ID, which may have biased RI estimates and

contributed to the observed heterogeneity. Lower RI limits from many studies differed from the WHO's age-specific thresholds for defining ID.[16] While RIs and WHO thresholds are defined differently and are expected to differ, they are both used in the definition and treatment of ID. For example, in children aged 24-33 months, the pooled estimates of the lower RI limit from the systematic review was 6-7 μg/L, whereas WHO recommends a threshold for ID of 12 μg/L.

We concluded the paper by calling for the development of rigorous, locally derived RIs aligned with known risk factors. Furthermore, RCs employing continuous age modeling were recommended as a means to mitigate the limitations of fixed age partitions and the substantial inconsistencies that contributed significantly to the heterogeneity across the RI estimates.

7.3 Summary of manuscript 4 (Chapter 5): Beyond reference intervals: optimizing lower and upper limits for reporting normative hemoglobin levels for children

In this paper, we presented novel RCs and OCs for hemoglobin. Traditional methods for interpreting hemoglobin levels based on RIs may lack the clinical granularity needed for detailed decision-making.[20,23] In this study, we moved beyond fixed age partitioned RIs and developed sex-specific reference and "optimal" curves. RCs overcome common sample size related limitations and allow RI estimates to be provided for any age or age interval. OCs may more accurately define hemoglobin distributions in optimally growing healthy children by excluding individuals with physiological, environmental, or clinical risk factors for ID.[24,25]

Using a cross-sectional sample from a cohort of healthy Canadian children, we estimated RCs. Based on previous research[26–29] and recommendations[13,15,30] on factors associated with iron status we also created OCs, using pre-specified optimality criteria. We used nonparametric quantile regression models with restricted cubic splines to estimate both RCs and OCs (2.5th and

97.5th percentile curves with respect to age) for hemoglobin.[31–33] RIs and OIs for specific age or age group were then generated from RCs and OCs. The OCs were compared to RCs and WHO anemia thresholds.[19] Reference limits for relatively narrower age intervals and an interactive web-based tool were provided. Among this study's key findings, we identified that the lower limits for the OCs were consistently higher than the lower limits for the RCs for females across all ages and for males under two years of age. Differences in upper limits between RCs and OCs were negligible. When compared to WHO anemia thresholds,[16] the lower OC limits (2.5th percentile of the optimal population) was always lower, which is expected since WHO threshold is based on the 5th percentile of the distribution. Nevertheless, WHO thresholds were often higher than 5th percentile for children who met all optimality criteria. The difference is important given that both estimates were derived from the 5th percentile, indicating the need for revisiting currently recommended hemoglobin thresholds for these ages.

7.4 Summary of manuscript 5 (Chapter 6): Ferritin reference and optimal curves compared with thresholds in pre-adolescents

In this paper, reference and optimal curves for ferritin are presented. Similar to hemoglobin, interpretation of ferritin in the pediatric population is complicated by the dynamic developmental changes observed, especially in early childhood, with traditional ferritin RIs potentially failing to reflect these dynamic changes.[34] In this paper, we developed ferritin RCs, treating age as a continuous variable, and OCs, where measurements for children without physiological, environmental, or clinical risk factors for ID were used.

We used data from the same cohort as for manuscript 4 to estimate RCs for children aged 2 weeks to 10 years. Based on prior evidence[26–29] and recommendations on determinants of iron status[13,15,30], to estimate OCs we applied optimality criteria excluding children with ID, elevated

C-reactive protein, prematurity, low birth weight, underweight or overweight status, and low household income. Both reference and optimal ferritin limits were markedly elevated in early infancy, followed by a steep decline during the first year of life, aligned with our systematic review findings from manuscript 3. Throughout infancy and early childhood, lower limits derived from OCs were consistently higher than those from RCs, reflecting the exclusion of children with clinical, environmental, or socioeconomic risk factors for ID. These differences narrowed with age, suggesting that such risk factors have the greatest influence when ferritin stores are the most variable. The 5th, 10th, and 20th percentiles of ferritin derived from the population with optimal iron status can be used to reveal differences in classification of iron status based on thresholds used. For instance, the WHO threshold for ferritin (<12 μg/L)[16] typically aligned between the 5th and 10th percentiles, while ASH threshold for ferritin (<20 μg/L)[35] often exceeded the 20th percentile, particularly after nine months of age, suggesting a potential for overdiagnosis. At twelve months, for example, ten percent of children meeting all optimality criteria would be labeled iron deficient using the WHO threshold, and more than twenty percent using the Ontario threshold.

7.5 Main points of integration

7.5.1 Current practice related to pediatric biomarkers of iron status

My dissertation examined current practices related to the development of pediatric RIs and RCs for hemoglobin and ferritin through two systematic reviews. In both systematic reviews, substantial heterogeneity was identified in published pediatric RIs and RCs for both biomarkers. Studies varied extensively in age partitioning, inclusion criteria, analyzer type, and sample size. Few studies followed CLSI guidelines and no study excluded children with anemia or risk factors for good iron status. Pooled estimates were only possible for selected age groups due to

inconsistent utilization of age intervals and sparse data. These limitations underline the need for standardized approaches and international harmonization.

The study by Braat et al. (2024), commissioned by WHO, pooled international databases.[19] The authors estimated hemoglobin thresholds at the 5th percentile for discrete age groups. The study by Pasricha et al. (2024) argued that this is "an opportunity for global harmonization of hemoglobin thresholds to define anemia across countries, clinical guidelines, and diagnostic laboratories".[36] Nevertheless, a key finding in both of our systematic reviews[21] and through detail exploration of our Canadian estimates was the discrepancy between locally derived RIs, the new WHO thresholds, and emerging CDLs. For ferritin, we similarly identified that lower RIs in early childhood were often noticeably lower than WHO thresholds and extensively lower than emerging CDLs reported in recent studies. This underscores the need to reconcile the value of local versus global thresholds.

In particular, the evidence behind current ferritin cut-offs to define ID in children is extremely limited. A recent Cochrane review of the diagnostic accuracy of ferritin compared with bone marrow iron identified one study only conducted in Malawian children, a small study done by Jonker et al. (2014). Both the WHO and the ASH Guideline Panel's draft recommendation for pediatric ferritin thresholds are based in part on the findings from this single study, rather than use of ferritin distributions such as the 5th percentile or lower limit of reference or optimal intervals and curves. Two studies by Mei and colleagues (2021, 2023),[37,38] using multiple NHANES national surveys, developed ferritin thresholds by examining the point at which hemoglobin begins to decline and soluble transferrin receptor begins to rise. They identified a physiological inflection point of approximately 20 µg/L for young children, which is substantially higher than WHO threshold of 12 µg/L but aligned with the proposed new ASH

guideline. Abdullah et al. (2017)[39] similarly evaluated ferritin cut-offs in relation to hemoglobin among 1-3 year old children using data from the TARGet Kids! cohort. Applying restricted cubic spline models to 1,257 healthy preschoolers, the authors demonstrated a non-linear relationship between hemoglobin and ferritin, with hemoglobin increasing steeply until a plateau was reached at ferritin concentrations around 18-24 µg/L and then rising only minimally thereafter. Furthermore, a recent modified Delphi study by Naveed et al. (2023)[40] reached expert consensus that currently used ferritin cut-offs are set too low and contribute to substantial underdiagnosis of ID across clinical settings. The study supported raising ferritin CDLs to below 30 µg/L in adults. Tang and colleagues (2025)[41] have launched the "Raise the Bar" initiative, a targeted advocacy effort encouraging laboratories and clinicians to adopt higher ferritin thresholds. In the 2025 draft recommendations, the ASH Guideline Panel formally suggested that clinicians use a pediatric ferritin threshold of ≤ 20 µg/L for diagnosis of ID in children aged 9 months to 4 years.

Thus, while there is ongoing controversy about a lower serum ferritin threshold to define ID in children, most expert groups are converging on a cut-off that is much higher than the lower reference limit (2.5th percentile) from available RIs/RCs, including those we established. Jäger et al. (2024) demonstrated that the choice of a threshold is consequential as the prevalence of ID changes sharply depending on whether clinicians use a ferritin cut-off of 15 µg/L or 30 µg/L.[42] Their work shows that controversies around ferritin thresholds have real clinical consequences, and this dissertation mirrors this pattern in pediatrics.

7.5.2 Beyond traditional approaches for developing pediatric reference intervals

Traditional RIs in pediatrics are often partitioned by age and sex, but this discrete approach does not fully capture the continuous biological variation that occurs throughout development.[20,23] For

rapidly changing biomarkers such as hemoglobin and ferritin, discrete partitions can create artificial boundaries and lead to misclassification, particularly for children who fall near the boundaries of an age partition.[20,34] The RCs provide a more robust alternative by modeling age as a continuous variable. In this way, RCs avoid abrupt shifts in interpretation when children transition between partitions.[20] This mirrors the use of growth centile curves in pediatric practice, which are already widely accepted for height and weight.[24,25] Moreover, use of RCs eliminates excessive partitioning, hence overcomes the common problem in the pediatric populations related to acquiring adequate samples. Furthermore, OCs extend the concept of RCs by further restricting the healthy reference population to those with characteristics associated with optimal iron status. In this dissertation, for both hemoglobin and ferritin, OCs based on pre-specified optimality criteria excluding children with factors unfavourable to good iron status were consistently higher than RCs, though differences were small. The small differences we observed likely reflects the underlying healthy status of children participating in the TARGet Kids! cohort, which may already have been favourable to good iron status. This finding translates the rationale behind WHO growth standards that in population where children "grow" under near optimal conditions, the observed reference curve may closely approach the optimal curve of how children "should grow".[24,25]

Both RCs and OCs provide information about the distribution of biomarkers in a reference population, against which international thresholds, such as WHO and those proposed by ASH, can be contextualized. RCs show how many children in a healthy reference population would be classified as iron deficient or anemic under a fixed international threshold at each age. OCs may help researchers and clinicians to better understand whether those fixed thresholds are appropriate for the local population. Together, these curves do not validate international

thresholds but rather help clarify how these fixed thresholds intersect with age dependent biomarker patterns. For ferritin, for example, a threshold of <20 μg/L proposed by the ASH classifies substantially more young children as having ID than either the 2.5th percentile of the OC or WHO's threshold of 12 μg/L. For hemoglobin, WHO 2024 anemia thresholds were frequently higher than even the 5th percentile of the optimal distribution in certain age groups. Our findings also illustrated that the proportion of children classified as having ID based on a fixed threshold may vary substantially by age. For example, using the ASH ferritin cut-off at 18 months of age, 55% of males and 37% of females in our cohort would be considered iron deficient, a prevalence far higher than what would be at 9 months or after 3.5 years.

7.5.3 Computational and methodological frameworks

A central challenge across both of the systematic reviews I conducted was the substantial inconsistency in age partitioning across the studies, which fundamentally limited the researcher's ability to synthesize evidence on pediatric RIs. Hemoglobin and ferritin vary continuously with age, yet published studies used age categories that ranged from 48-hour neonatal windows to age intervals spanning more than a decade. This huge difference in age intervals used and the significant heterogeneity it induces prevented direct pooling in many cases.

To address the limitation of inconsistency in age partitioning, this dissertation introduced a standardized age partitioning framework: 3-month intervals from birth to 3 years and 1-year intervals thereafter. This structure allowed heterogeneous studies to be mapped onto a standardized age partitions, enabling quantitative synthesis and improving narrative synthesis where pooling was not possible. Standardized partitioning also supports the PRINCES-H and PRINCES-FERRITIN Shiny applications developed as part of this dissertation. These interactive tools support harmonization by allowing users to overlay dozens of RI and RC studies, examine

sex-specific distributions, and compare RI and RC results against WHO thresholds and the ASH draft recommendations. Because the apps dynamically reorganize evidence into standardized age partitions, they support real time synthesis of complex data from multiple studies that is not feasible with static tables or figures. This approach not only mitigates a major obstacle in evidence synthesis involving RIs and RCs but also creates a framework and infrastructure for future updates, allowing new studies or CDLs to be integrated seamlessly into living RIs and OIs as well as RCs and OCs.

The integrated framework presented here combining evidence synthesis, harmonized age partitioning, robust curve estimating approaches, and introduction of OCs and OIs offers several contributions. First, it provides a methodological roadmap for synthesizing heterogeneous literature on pediatric RIs and RCs, including tools for visualizing global variation (PRINCES-H and PRINCES-FERRITIN), which can be tailored and adapted for any pediatric biomarker. Second, it demonstrates the feasibility and value of curve based approaches that treat age as continuous, improving precision while reducing arbitrary age partitioning and optimizing available sample size. Third, it lays groundwork for use of optimality criteria in pediatric biomarker research, extending the conceptual logic of WHO growth standards to iron related biomarkers and laboratory biomarkers in general. Fourth, by developing interactive web applications, this dissertation establishes a "living reference and optimal interval and curve" platform that can integrate new data, compare distributions to existing (WHO or ASH) and emerging thresholds, and support both clinical interpretation and health assessment of iron status of a population.

7.6 Strengths and limitations

A major strength of this dissertation is that this is the first comprehensive systematic review and meta-analysis that allowed for synthesizing, in manuscripts 2 and 3, existing literature on pediatric RIs and RCs for hemoglobin and ferritin using standardized age partitioning. We concluded that robust RCs that treat age as continuous variable were needed, which was another contribution of this dissertation through manuscripts 2 and 3. An additional strength was the use of a large-scale dataset (TARGet Kids! cohort study)[43] for RC estimation that included children as young as 2-weeks of age, which allowed estimation of both reference and optimal curves for pre-adolescent's population in manuscripts 4 and 5. I was also able to rely on cohort data documenting risk factors for good irons status to estimate OCs, which was a novel contribution of this research. Finally, an overarching strength across all 4 manuscripts was the development of graphical interactive tools that allow for complex visualization of the findings: heterogeneity across published literature, narrative synthesis by custom age-specific intervals, and direct comparison with existing or emerging international thresholds and individualized patient's lab results.

There are several limitations in manuscripts 2 and 3. One such limitation was that we did not conduct formal risk of bias or quality assessment for studies estimating RIs/RCs. We used adherence to the CLSI guideline as an indicator of quality given the lack of appropriate established risk of bias tools for these studies. We also included only English-language publications in the reviews, which might have led to missing important studies published in other languages. We also excluded non-English speaking families from the cohort study-based analyses presented in manuscripts 4 and 5, which was an exclusion made by the cohort study researchers that may limit the generalizability of our findings.

7.7 Future research directions

Findings of this dissertation emphasize the critical importance of ongoing and future pediatric data collection efforts, including gathering blood samples together with health information from healthy children, especially for children under 2-weeks of age and over 11-years. The dissertation presented work as proof-of-concept for future studies in the field of laboratory medicine, including the methods employed and the analysis strategies used, and my novel computational and graphical web-based applications. This work can be adapted and used in evidence synthesis of RIs and RCs and estimation of RCs and OCs for other pediatric biomarkers. Future studies should utilize knowledge translation, engaging laboratories, clinicians, and public health practitioners to understand the feasibility, usability, and clinical value of these web-based applications. And lastly, future collaborative efforts should be made for future studies to assemble and integrate international pediatric datasets to derive multi country RCs and OCs estimates and examine the impact of regional factors on these curves.

7.8 Overall conclusion

This dissertation identifies key gaps and limitations in the current landscape of pediatric RIs for hemoglobin and ferritin and their interpretation, presents new RCs and OCs based on robust data from Canadian children, and proposes frameworks for generating future reliable RIs, OIs, RCs, and OCs. The integrated findings support three overarching conclusions: (1) existing pediatric RIs and RCs are highly heterogeneous and often misaligned with global thresholds for identifying and defining ID; (2) curve-based estimation can improve precision and interpretability, especially during early childhood when biomarker distributions change rapidly, and (3) even as single thresholds gain prominence, understanding age- and sex-specific reference and optimal distributions remain essential for accurate diagnosis, health surveillance, and

evaluation of emerging clinical guidelines. This research can be extended to other pediatric biomarkers and eventually to other populations.

References

1 Ceriotti F, Hinzmann R, Panteghini M. Reference intervals: the way forward. *Ann Clin Biochem Int J Lab Med* 2009; **46**: 8–17.

2 Clinical and Laboratory Standards Institute (CLSI), editor. Defining, Establishing and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline—Third Edition. Wayne, Pa: Clinical and Laboratory Standards Institute, 2008.

3 Ozarda Y. Reference intervals: current status, recent developments and future considerations. *Biochem Medica* 2016; : 5–16.

4 Adeli K, Higgins V, Trajcevski K, White-Al Habeeb N. The Canadian laboratory initiative on pediatric reference intervals: A CALIPER white paper. *Crit Rev Clin Lab Sci* 2017; **54**: 358–413.

5 Daly CH, Liu X, Grey VL, Hamid JS. A systematic review of statistical methods used in constructing pediatric reference intervals. *Clin Biochem* 2013; **46**: 1220–7.

6 Zierk J, Baum H, Bertram A, *et al.* High-resolution pediatric reference intervals for 15 biochemical analytes described using fractional polynomials. *Clin Chem Lab Med CCLM* 2021; **59**: 1267–78.

7 Brouwer MJ, De Vries LS, Groenendaal F, *et al.* New Reference Values for the Neonatal Cerebral Ventricles. *Radiology* 2012; **262**: 224–33.

8 Faulkner WR, Demers LM. Importance of age-dependent reference values in biochemical testing: are we including the elderly population? *Clin Chem* 1994; **40**: 855–6.

9 Rosner B, Cook N, Portman R, Daniels S, Falkner B. Determination of Blood Pressure Percentiles in Normal-Weight Children: Some Methodological Issues. *Am J Epidemiol* 2008; **167**: 653–66.

10 Adeli K, Higgins V, Seccombe D, *et al.* National Survey of Adult and Pediatric Reference Intervals in Clinical Laboratories across Canada: A Report of the CSCC Working Group on Reference Interval Harmonization. *Clin Biochem* 2017; **50**: 925–35.

11 Higgins V. What is the current state of reference intervals in Canada? A need for harmonization. 2017; published online Nov 20. https://myadlm.org/science-and-research/scientific-shorts/2017/current-state-of-reference-intervals-in-canada.

12 George-Gay B, Parker K. Understanding the complete blood count with differential. *J Perianesth Nurs* 2003; **18**: 96–117.

13 Baker RD, Greer FR, The Committee on Nutrition. Diagnosis and Prevention of Iron Deficiency and Iron-Deficiency Anemia in Infants and Young Children (0–3 Years of Age). *Pediatrics* 2010; **126**: 1040–50.

14 Christofides A, Schauer C, Zlotkin SH. Iron deficiency anemia among children: Addressing a global public health problem within a Canadian context. *Paediatr Child Health* 2005; **10**: 597–601.

15 Unger SL, Fenton TR, Jetty R, Critch JN, O'connor DL. Iron requirements in the first 2 years of life. *Paediatr Child Health* 2019; **24**: 555–555.

16 World Health Organization. WHO guideline on use of ferritin concentrations to assess iron status in individuals and populations. Geneva: World Health Organization, 2020 https://iris.who.int/handle/10665/331505 (accessed June 18, 2025).

17 Jorgensen JM, Crespo-Bellido M, Dewey KG. Variation in hemoglobin across the life cycle and between males and females. *Ann N Y Acad Sci* 2019; : nyas.14096.

18 World Health Organization, editor. Guideline on haemoglobin cutoffs to define anaemia in individuals and populations. Geneva: World Health Organization, 2024.

19 Braat S, Fielding KL, Han J, *et al.* Haemoglobin thresholds to define anaemia from age 6 months to 65 years: estimates from international data sources. *Lancet Haematol* 2024; **11**: e253–64.

20 Daly CH, Higgins V, Adeli K, Grey VL, Hamid JS. Reference interval estimation: Methodological comparison using extensive simulations and empirical data. *Clin Biochem* 2017; **50**: 1145–58.

21 Bijelić V, Bijelić M, Larock J, *et al.* Paediatric Reference Intervals and Curves for Haemoglobin Estimated Using Direct Methods: A Systematic Review and Meta-Analysis. *Int J Lab Hematol* 2025; : ijlh.14489.

22 Bijelić V, Potter B, Parkin PC, Momoli F, Liebman M, Hamid JS. Paediatric reference intervals and curves for haemoglobin and ferritin: protocol for a systematic review and meta-analysis. *BMJ Open* 2023; **13**: e073783.

23 Higgins V, Adeli K. Advances in pediatric reference intervals: from discrete to continuous. *J Lab Precis Med* 2018; **3**: 3–3.

24 De Onis M, Garza C, Habicht J-P. Time for a New Growth Reference. *Pediatrics* 1997; **100**: e8–e8.

25 de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J. The WHO Multicentre Growth Reference Study: planning, study design, and methodology. *Food Nutr Bull* 2004; **25**: S15-26.

26 Bayoumi I, Parkin PC, Birken CS, Maguire JL, Borkhoff CM, for the TARGet Kids! Collaboration. Association of Family Income and Risk of Food Insecurity With Iron Status in Young Children. *JAMA Netw Open* 2020; **3**: e208603.

27 Borkhoff SA, Parkin PC, Birken CS, Maguire JL, Macarthur C, Borkhoff CM. Examining the Double Burden of Underweight, Overweight/Obesity and Iron Deficiency among Young Children in a Canadian Primary Care Setting. *Nutrients* 2023; **15**: 3635.

28 Oatley H, Borkhoff CM, Chen S, *et al.* Screening for Iron Deficiency in Early Childhood Using Serum Ferritin in the Primary Care Setting. *Pediatrics* 2018; **142**: e20182095.

29 Sypes EE, Parkin PC, Birken CS, *et al.* Higher Body Mass Index Is Associated with Iron Deficiency in Children 1 to 3 Years of Age. *J Pediatr* 2019; **207**: 198-204.e1.

30 Domellöf M, Braegger C, Campoy C, *et al.* Iron Requirements of Infants and Toddlers. *J Pediatr Gastroenterol Nutr* 2014; **58**: 119–29.

31 Van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med* 1999; **18**: 681–94.

32 Koenker R. quantreg: Quantile Regression. 2024. https://CRAN.R-project.org/package=quantreg.

33 Harrell Jr FE. rms: Regression Modeling Strategies. 2024. https://CRAN.R-project.org/package=rms.

34 Lyle AN, Pokuah F, Dietzen DJ, *et al.* Current State of Pediatric Reference Intervals and the Importance of Correctly Describing the Biochemistry of Child Development: A Review. *JAMA Pediatr* 2022; **176**: 699.

35 Wozniak, M., Catomeris, P., Li, H., *et al.* Guidelines for the Use of Laboratory Tests for Iron Deficiency (CLP 002). Ontario Association of Medical Laboratories, 2024.

36 Pasricha S-R, Rogers L, Branca F, Garcia-Casal M-N. Measuring haemoglobin concentration to define anaemia: WHO guidelines. *The Lancet* 2024; **403**: 1963–6.

37 Mei Z, Addo OY, Jefferds ME, Sharma AJ, Flores-Ayala RC, Brittenham GM. Physiologically based serum ferritin thresholds for iron deficiency in children and non-pregnant women: a US National Health and Nutrition Examination Surveys (NHANES) serial cross-sectional study. *Lancet Haematol* 2021; **8**: e572–82.

38 Mei Z, Addo OY, Jefferds MED, *et al.* Comparison of Current World Health Organization Guidelines with Physiologically Based Serum Ferritin Thresholds for Iron Deficiency in Healthy Young Children and Nonpregnant Women Using Data from the Third National Health and Nutrition Examination Survey. *J Nutr* 2023; **153**: 771–80.

39 Abdullah K, Birken CS, Maguire JL, *et al.* Re-Evaluation of Serum Ferritin Cut-Off Values for the Diagnosis of Iron Deficiency in Children Aged 12-36 Months. *J Pediatr* 2017; **188**: 287–90.

40 Naveed K, Goldberg N, Shore E, *et al.* Defining ferritin clinical decision limits to improve diagnosis and treatment of iron deficiency: A modified Delphi study. *Int J Lab Hematol* 2023; **45**: 377–86.

41 Tang GH, Phillips Z, Poutanen S, *et al.* 'Raise the Bar', a Canadian intervention to combat disparities in the recognition and management of iron deficiency: Development of the intervention, feasibility and implementation. *Br J Haematol* 2025; **207**: 2219–23.

42 Jäger L, Rachamin Y, Senn O, Burgstaller JM, Rosemann T, Markun S. Ferritin Cutoffs and Diagnosis of Iron Deficiency in Primary Care. *JAMA Netw Open* 2024; **7**: e2425692.

43 Carsley S, Borkhoff CM, Maguire JL, *et al.* Cohort Profile: The Applied Research Group for Kids (TARGet Kids!). *Int J Epidemiol* 2015; **44**: 776–88.