

Haemoglobin thresholds to define anaemia from age 6 months to 65 years: estimates from international data sources

Sabine Braat*, Katherine L Fielding*, Jiru Han, Victoria E Jackson, Sophie Zaloumis, Jessica Xu Hui Xu, Gemma Moir-Meyer, Sophia M Blaauwendraad, Vincent W V Jaddoe, Romy Gaillard, Patricia C Parkin, Cornelia M Borkhoff, Charles D G Keown-Stoneman, Catherine S Birken, Jonathon L Maguire, Genes & Health Research Team,† Melanie Bahlo, Eliza M Davidson, Sant-Rayn Pasricha



Summary

Background Detection of anaemia is crucial for clinical medicine and public health. Current WHO anaemia definitions are based on statistical thresholds (fifth centiles) set more than 50 years ago. We sought to establish evidence for the statistical haemoglobin thresholds for anaemia that can be applied globally and inform WHO and clinical guidelines.

Methods In this analysis we identified international data sources from populations in the USA, England, Australia, China, the Netherlands, Canada, Ecuador, and Bangladesh with sufficient clinical and laboratory information collected between 1998 and 2020 to obtain a healthy reference sample. Individuals with clinical or biochemical evidence of a condition that could reduce haemoglobin concentrations were excluded. We estimated haemoglobin thresholds (ie, 5th centiles) for children aged 6–23 months, 24–59 months, 5–11 years, and 12–17 years, and adults aged 18–65 years (including during pregnancy) for individual datasets and pooled across data sources. We also collated findings from three large-scale genetic studies to summarise genetic variants affecting haemoglobin concentrations in different ancestral populations.

Findings We identified eight data sources comprising 18 individual datasets that were eligible for inclusion in the analysis. In pooled analyses, the haemoglobin fifth centile was 104.4 g/L (90% CI 103.5–105.3) in 924 children aged 6–23 months, 110.2 g/L (109.5–110.9) in 1874 children aged 24–59 months, and 114.4 g/L (113.6–115.2) in 1839 children aged 5–11 years. Values diverged by sex in adolescents and adults. In pooled analyses, the fifth centile was 122.2 g/L (90% CI 121.3–123.1) in 1741 female adolescents aged 12–17 years and 128.2 g/L (126.4–130.0) in 1103 male adolescents aged 12–17 years. In pooled analyses of adults aged 18–65 years, the fifth centile was 119.7 g/L (90% CI 119.1–120.3) in 3640 non-pregnant females and 134.9 g/L (134.2–135.6) in 2377 males. Fifth centiles in pregnancy were 110.3 g/L (90% CI 109.5–111.0) in the first trimester (n=772) and 105.9 g/L (104.0–107.7) in the second trimester (n=111), with insufficient data for analysis in the third trimester. There were insufficient data for adults older than 65 years. We did not identify ancestry-specific high prevalence of non-clinically relevant genetic variants that influence haemoglobin concentrations.

Interpretation Our results enable global harmonisation of clinical and public health haemoglobin thresholds for diagnosis of anaemia. Haemoglobin thresholds are similar between sexes until adolescence, after which males have higher thresholds than females. We did not find any evidence that thresholds should differ between people of differencing ancestries.

Funding World Health Organization and the Bill & Melinda Gates Foundation.

Copyright © 2024 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Introduction

Anaemia develops when the red cell mass is insufficient to meet physiological oxygen-carrying needs and is usually formally diagnosed when the haemoglobin concentration falls below a defined threshold for age and sex.¹ Accurate case definition of anaemia is crucial for clinical diagnosis and treatment, and for understanding the magnitude and distribution of this condition as a public health problem. WHO's anaemia definitions aim to reflect the lower fifth centile of the haemoglobin distribution of a reference population of healthy individuals. WHO thresholds were initially proposed in 1958,² then updated

in 1968,³ and have remained essentially unchanged since.⁴ WHO thresholds were based on studies with restricted measurement of biomarkers of iron and other forms of haematinic deficiency and inflammation. There is limited consensus on anaemia definitions,⁵ leading to heterogeneous definitions between sources, expert groups, and public health bodies,⁶ translating into inconsistent clinical definitions.⁷

Establishing a valid diagnosis of anaemia is crucial for treating patients and detecting the range of diseases that could underlie this condition.⁸ Anaemia can cause clinical effects directly (eg, anaemia causes a substantial burden

Lancet Haematology 2024; 11: e253–64

Published Online
February 29, 2024
[https://doi.org/10.1016/S2352-3026\(24\)00030-9](https://doi.org/10.1016/S2352-3026(24)00030-9)
See [Comment](#) page e242

*Contributed equally

†Members of the Genes & Health Research Team are listed in the appendix (p 195)

Population Health and Immunity Division, The Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia (S Braat MSc, K L Fielding MBBS, J Han PhD, V E Jackson PhD, G Moir-Meyer PhD, Prof M Bahlo PhD, E M Davidson PhD, Prof S-R Pasricha PhD);

Methods and Implementation Support for Clinical and Health research Hub (S Braat, S Zaloumis PhD, J X H Xu MBIostat), and Medical Biology (K L Fielding, J Han, V E Jackson, G Moir-Meyer, Prof M Bahlo, Prof S-R Pasricha), Faculty of Medicine, Dentistry and Health Sciences, The University of Melbourne, Melbourne, VIC, Australia; Clinical Haematology, The Austin Hospital, Heidelberg, VIC, Australia (K L Fielding); Generation R Study Group, and Department of Pediatrics, Erasmus University Medical Center, Rotterdam, Netherlands

(S M Blaauwendraad MD, Prof V W V Jaddoe PhD, R Gaillard PhD); Division of Pediatric Medicine and the Pediatric Outcomes Research Team, The Hospital for Sick Children, Toronto, ON, Canada (P C Parkin MD, C M Borkhoff PhD, Prof C S Birken MD); Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada (P C Parkin, C M Borkhoff,

C D G Keown-Stoneman PhD, Prof C S Birken, J L Maguire MD); Unity Health Toronto, Toronto, ON, Canada (C D G Keown-Stoneman, J L Maguire); **Wolfson Institute of Population Health and Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK** (Genes & Health Research Team); **Diagnostic Haematology, The Royal Melbourne Hospital, Parkville, VIC, Australia** (Prof S-R Pasricha); **Clinical Haematology, Peter MacCallum Cancer Centre and The Royal Melbourne Hospital, Parkville, VIC, Australia** (Prof S-R Pasricha)

Correspondence to: Professor Sant-Rayn Pasricha, Population Health and Immunity Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
pasricha.s@wehi.edu.au
See [Online](#) for appendix

Research in context

Evidence before this study

WHO's haemoglobin thresholds to define anaemia are statistical thresholds reflecting the fifth centile of haemoglobin distribution in apparently healthy populations, set more than 50 years ago and maintained since then. Before undertaking the current analysis, we first evaluated available documentation of the source of these recommendations: the thresholds were defined on the basis of five small studies, which were restricted in terms of geographical location, epidemiological design, sample size, and laboratory evaluation of iron stores and inflammation as well as other haematinic biomarkers, and transparency of the statistical analysis. These studies included minimal data from women and did not include data from children or infants. Next, we searched for studies that had sought to evaluate thresholds, using a 2019 WHO-commissioned review as an initial resource. This review summarised results from 60 studies conducted globally between 1975 and 2018, and noted markedly heterogeneous results, with studies inconsistently excluding individuals with iron deficiency or inflammation, and a recommendation to conduct studies with more rigorous inclusion criteria. More recently, a post-hoc analysis of international datasets from population nutrition surveys through the BRINDA (Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia) Collaboration illustrated the constraints of the use of population haematology (collected with a mixture of capillary and venous sampling) and biochemical data with insufficient clinical information, showing an approximately 33 g/L variation in fifth centile of haemoglobin concentration across different countries (eg, finding a threshold for young children aged 6–59 months ranging from approximately 79 g/L in Pakistan and Cote D'Ivoire to approximately 113 g/L in the USA; and a threshold in

non-pregnant women ranging from 88 g/L in India to 121 g/L in the USA). An analysis based on data from a national nutrition survey conducted in India also suggested thresholds lower than WHO recommendations in children and adolescents.

Added value of this study

In this analysis, we identified and used datasets that enabled exclusion of clinical conditions that could influence haemoglobin concentrations, along with biochemical measures of iron status, inflammation, and, where possible, renal and other organ function, to create reference populations of healthy individuals from which the fifth centile of the haemoglobin distribution could be estimated. Our data include children aged 6 months to 18 years, and adults aged up to 65 years. Our results suggest relatively consistent haemoglobin thresholds across diverse populations, including from south and east Asia, and reveal minimal secular variation within countries even as populations (and reasons for exclusion) changed over recent decades. We also summarised findings from three existing large-scale, multi-ancestry, genome-wide association studies to search for non-clinically relevant, high-frequency genetic variants that might influence haemoglobin concentrations in populations of different ancestries, and that could hence justify varied haemoglobin thresholds between different ancestries.

Implications of all the available evidence

Haemoglobin thresholds defining anaemia in representative healthy individuals vary across the lifecycle and differ by sex in people aged 12–65 years. Additionally, our analyses of genetic variants indicate that haemoglobin thresholds for anaemia should not be adjusted for individuals of different ancestry. Our analyses provide a single set of thresholds that can be applied globally in clinical practice and public health.

of disease); however, even in the absence of overt clinical effects, haemoglobin-based thresholds for anaemia are used as important triggers for investigation and optimisation of underlying contributors. Haemoglobin concentrations are reduced by acute, recurrent, chronic medical or surgical illness. Inflammation (eg, due to autoimmune disease, infection, cancer, heart failure, and even obesity) suppresses erythropoiesis due to functional iron deficiency and might also reduce red cell survival.⁸ Haemoglobin concentrations are reduced in renal impairment due to reduced erythropoietin production and functional iron deficiency. Many medications can reduce haemoglobin concentrations idiosyncratically or in a dose-dependent manner. Anaemia can persist weeks or months beyond an acute illness^{9,10} and normalisation of acute inflammatory markers. Conversely, hypoxia (eg, smoking, respiratory or cardiac disease, obesity or sleep apnoea, or elevated altitude of residence)¹¹ can increase haemoglobin concentrations. Diagnosis of anaemia lies at the centre of many clinical pathways, such as investigation of gastrointestinal bleeding,¹² preoperative

optimisation,¹³ screening for haemoglobinopathy,¹⁴ and eligibility for blood donation.¹⁵ Anaemia is an adverse prognostic factor for multiple conditions, including heart failure,¹⁶ major surgery,¹⁷ cancer,¹⁸ and HIV.¹⁹ Most patients admitted to hospital with critical illness are anaemic in hospital, and about half remain anaemic 6 months²⁰ or even 12 months after hospital admission.⁹

Valid definitions of anaemia are also crucial for guiding population interventions and tracking progress towards global targets. Based on current definitions, anaemia affects 40% of all preschool children and 30% of women, including 46% of pregnant women globally,²¹ and accounts for 52 million years lived with disability.²² Reducing anaemia prevalence in women by 50% is a WHO global nutrition target²³ and a sub-indicator of the UN Sustainable Development Goals.²⁴ Interventions that directly (eg, iron supplementation²⁵ or fortification²⁶) or indirectly (eg, malaria control) reduce anaemia are recommended for many low-income countries, with implementation decisions based on anaemia prevalence.²⁷

Haemoglobin thresholds to define anaemia usually vary between men and women, although this distinction has been challenged.^{28,29} Thresholds can also vary in children and during pregnancy. Haemoglobin concentrations can be lower in African^{5,30} or Asian populations,³¹ but it is unclear whether this discrepancy reflects clinically significant genetic conditions affecting red cells, undefined clinical illnesses, or a different underlying baseline driven by non-clinical genetic variation.³⁰

WHO has recognised the need for further evidence for haemoglobin thresholds to define anaemia, including the influence of genetic factors (both carrier states for rare genetic conditions such as thalassaemias and haemoglobinopathies, and the role of genetics in overall variation),³² in order to update global guidelines. Here, we present an analysis utilising multiple large-scale datasets in ethnically diverse populations that included sufficiently detailed clinical and laboratory parameters to enable post-hoc assembly of reference populations of healthy individuals without discernible risk factors for anaemia, in which the fifth centile of haemoglobin distribution can be estimated across the lifecycle. Additionally, we summarised ancestry-specific allele frequencies and effects on haemoglobin concentrations of genetic polymorphisms and structural variants identified through existing large-scale, multi-ancestry genome-wide association studies (GWAS). We aimed to provide a rationale for a set of thresholds that can be applied globally for clinical and population use.

Methods

Study design and data sources

To support clinical and public health guidelines we sought to establish thresholds for children, adolescents, adult men, and adult women including pregnant women (by trimester). We aimed to define the lower fifth centile of the distribution among healthy individuals (ie, a healthy reference population). Central to the design of the analysis was the need to establish a reference population from existing datasets (posteriori approach) through exclusion of individuals with evidence of conditions or circumstances that could influence haemoglobin concentration.³³ We searched for datasets comprising appropriate clinical and laboratory information to enable post-hoc derivation of an apparently healthy reference sample; this approach included assessing national health surveys (using the search terms “country name” AND “national health database”) and prospective clinical cohorts, with a focus on countries where malaria was not endemic (ie, with low background inflammation) between May, 2021, and September, 2022 (further details are provided in the appendix, p 5).

Although exclusion criteria were standardised, the available data and their coding varied between studies (eg, in classification of ancestry, consumption of alcohol, terminology, or questions about history of illness). Studies were only selected if both clinical and laboratory

data were available. The specific nature of clinical data varied between surveys, and within each survey all available clinical parameters relevant to anaemia were used. This approach resulted in heterogeneous exclusion criteria between surveys, but the optimal healthy population within each survey. Clinical criteria for exclusion included reporting of chronic systemic medical illness; any recent acute illness (2–4 weeks); recent hospital admission (up to 12 months); use of medications (excluding hormonal contraception where possible); current smoking or excess alcohol consumption; and obesity or low bodyweight (ascertained by BMI cutoffs). Where data were available, we excluded individuals living at elevations above 750 m. Missing clinical criteria were not applied as exclusions, except for pregnancy.

Laboratory criteria were required for all participants; those missing data for haemoglobin, ferritin, or C-reactive protein (CRP) were excluded from the analysis. Haemoglobin measurements were conducted on venous blood by use of an automated analyser or high-quality point-of-care device (table). Ferritin concentrations were applied as per WHO guidelines for the use of ferritin, defining iron deficiency as ferritin concentrations lower than 12 µg/L in children younger than 5 years, and ferritin concentrations lower than 15 µg/L in all other groups.³⁴ CRP higher than 5 mg/mL was used to indicate inflammation; individuals with elevated ferritin (as per WHO guidelines) were also excluded since this could portend inflammation or liver disease.³⁵ Where available, other markers of health were used on the basis of consensus reference ranges. Sensitivity analyses included raising the ferritin threshold to define iron deficiency or lowering the CRP threshold to define inflammation, or restricting mean cell volume (in children and adolescents only), or a combination of the above. Detailed inclusion and exclusion of individuals from each dataset to obtain the reference sample are shown in the appendix (pp 6–9).

We sought to ensure ethnic diversity but recognised that population field surveys in low-income settings can contain a high burden of unreported, undetected, or recently resolved inflammation even if acute biomarkers of inflammation had normalised, thus preventing exclusion of individuals with recent illness that might have lowered haemoglobin concentration unless detailed clinical data were included.³⁶ As recommended, we recalculated estimated glomerular filtration rates with non-race-based equations.³⁷

We sought to establish whether different ancestries have non-clinically relevant high-frequency genetic variations that influence haemoglobin concentration; the absence of such variants would enable a single set of thresholds. Thus, in a separate study to the determination of the fifth centile of the reference population described above, we used ancestry-specific databases to interrogate genome-wide associations, effect sizes, and minor allele frequencies of single-nucleotide polymorphisms (SNPs) and structural

	Design	Location	Study period	Population (sample size*; sample size final reference population)	Blood sample	Haemoglobin testing method
Australian Health Survey (AHS)	National population health survey	Australia	2011–12	Adults aged 18–65 years (N=7262; N=1268) Adolescents aged 12–17 years (N=734; N=380)	Non-fasting venous	Sysmex XE-2100
Benefits and Risks of Iron interventions in Children (BRISC)	Randomised controlled trial	Bangladesh	2017–20	Children aged 11 months (N=1365†; N=255)	Non-fasting venous	Hemocue 301+
China Health and Nutrition Survey (CHNS)	National population health survey	China‡	2009	Adults aged 18–65 years (N=7172; N=796) Children and adolescents aged 5–17 years (N=814; N=475)	Non-fasting venous	Sysmex XE-2100
Encuesta Nacional de Salud y Nutrición (ENSANUT)	National population health survey	Ecuador‡	2011–13	Children aged 6–59 months (N=2047; N=275)	Non-fasting venous	Sysmex XE-2100
Generation R	Prospective cohort study	Netherlands	2002–06	Pregnant women aged 18–45 years (N=7487; N=883)	Non-fasting venous	Automated Laboratory Analyzer§
Health Survey for England (HSE)	National population health survey	England	1998, 2006, 2009	Adults aged 18–65 years (N=15 315; N=2290)	Non-fasting venous	Coulter STKS, Abbott CD 4000
National Health and Nutrition Examination Survey (NHANES)	National population health survey	USA	1999–2000, 2001–02, 2003–04, 2005–06, 2007–08, 2009–10, 2015–16, 2017–18	Adults aged 18–65 years (N=19 346; N=1663) Children and adolescents aged 6 months to 17 years (N=12 883; N=4524)	Non-fasting venous	Beckman-Coulter analysers
TARGet Kids!	Open longitudinal cohort study of healthy children	Canada	2010–19	Children aged 6 months to 11 years (N=4700¶; N=1572)	Non-fasting venous	Sysmex XN-9000

CRP=C-reactive protein. The Australian Health Survey consists of the National Health Survey (NHS) and National Nutrition and Physical Activity Survey (NNPAS). The BRISC trial is registered on the Australian New Zealand Clinical Trials Registry (ACTRN12617000660381). The Generation R study is registered with the Dutch Trial Register (NL6484). The TARGet Kids! study is registered with ClinicalTrials.gov (NCT01869530). All data sources included male and female participants (apart from Generation R, which only included pregnant women). *Defined as those within the age range of the populations (eg, aged 18–65 years) and with non-missing haemoglobin, ferritin, and CRP concentrations except for Generation R defined as those with non-missing haemoglobin. †Additionally restricted to those who received an iron intervention. ‡Exclusions applied to those who resided above 750 m altitude (ENSANUT) or residents of province (Guizhou) due to altitude above 750 m (CHNS). §Specific analyser unavailable; automated laboratory haematology instrument. ¶Number of unique participants.

Table: Summary of data sources and demographics

variants associated with haemoglobin concentration. Full details are available in the appendix (pp 22–23).

Each included study had received ethics approval and obtained informed consent from participants (appendix p 25); use of these datasets for these analyses was deemed by Walter and Eliza Hall Institute of Medical Research (WEHI) Governance, Risk and Compliance to meet the criteria for exemption from ethics review.

Statistical analysis

Lower 2.5th and fifth centiles were estimated and two-sided confidence intervals (CIs) obtained with the methodology described below and detailed in the appendix (pp 16–22). The Clinical & Laboratory Standards Institute recommends using a 90% CI.³⁸ We estimated discrete thresholds (based on conventionally applied age categories) for each data source and continuous thresholds (based on the participant's age [and sex if applicable] within age categories, except for pregnant women) across all data sources. For each threshold, the sample size was determined by the number of participants who fulfilled the criteria of an apparently healthy individual among those with a non-missing haemoglobin, ferritin, and CRP value. Outlying values were identified with Tukey's method,³⁹ and the results presented here are those after excluding outliers from the reference sample.

For discrete thresholds, the assumption of normality was checked and the parametric theoretical centile of a

Gaussian distribution, with corresponding 90% CI, was used to estimate the centile.⁴⁰ For the national population health surveys, we aimed to account for the survey design by estimating survey-weighted centiles calculated with survey-weighted quantile regression. The implementation of survey-weighted quantile regression proceeded as described for continuous thresholds below, except the model only included a single intercept term. Standard errors for the intercept term were derived from 1000 Canty and Davidson's bootstrap replicates and used to calculate 90% CIs assuming a normal approximation.^{41,42} As a sensitivity analysis, unweighted centile estimates were also calculated with the parametric theoretical centile described earlier. Discrete centile estimates were pooled across all data sources through a fixed-effects meta-analysis, with a random-effects meta-analysis as a sensitivity analysis, and presented in forest plots.

For continuous thresholds, we used the Hoq⁴³ method to estimate age-specific (and by sex, if applicable) continuous haemoglobin thresholds within each age category by combining individual data of all relevant data sources. This approach involved identifying the best fitting multivariable fractional polynomial model for the mean values of haemoglobin by use of Royston's method⁴⁴ followed by a likelihood ratio test for an interaction between sex and the fractional polynomial representation of age. An age-by-sex interaction term was included in the model if it was statistically significant at the nominal significance level ($p < 0.05$). Quantile regression was then used; no survey

weighting was applied to the national population health surveys. Age-dependent haemoglobin predictions were obtained from the unweighted quantile regression model and bootstrap percentile 90% CIs generated from 1000 bootstrap samples. Results were presented in plots to indicate how predicted percentiles change with age (and by sex, if applicable), superimposed onto a scatter plot of haemoglobin versus age within each age category.

Reference samples were derived with Stata (version 16.1), R (version 4.2.0, Generation R), or Stata (version 17.1) for datasets from Australia's National Health Survey (NHS) and Australia's National Nutrition and Physical Activity Survey (NNPAS). Haemoglobin thresholds were estimated with R (version 4.1.1 and version 4.2.0 [Generation R]), or Stata (version 17.1) for datasets from the NHS and NNPAS. Details of the packages used, including those for the genetic analysis, are provided in the appendix (pp 23–25).

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit the manuscript for publication.

Results

We identified eight data sources comprising 18 individual datasets that were eligible for inclusion in the analysis (appendix pp 10–16). Included studies are summarised in the table. Studies consisted of national health or nutrition surveys (Health Survey for England [HSE],⁴⁵ National Health and Nutrition Examination Survey [NHANES]⁴⁶ Australian Health Survey [AHS],⁴⁷ China Health and Nutrition Survey [CHNS], and Encuesta Nacional de Salud y Nutrición [ENSANUT]⁴⁸) observational studies (The Applied Research Group for Kids [TARGet Kids!],⁴⁹ and Generation R⁵⁰) and a clinical trial (Benefits and Risks of Iron interventions in Children [BRISC]).⁵¹ Information on sex or gender was self-reported, and data reported gender in NHANES, CHNS, ENSANUT, and Target Kids!, and sex in HSE, AHS, BRISC, and Generation R. All studies reported only male or female. In this report we refer to males and females using the

term sex. We separately analysed thresholds for adult men aged 18–65 years, adult women aged 18–65 years, children aged 6–59 months (further sub-categorised as 6–23 months and 24–59 months), children aged 5–11 years, male and female adolescents aged 12–17 years, and pregnant women aged 18–45 years (by trimester). Insufficient data were available for children younger

For more on the **China Health and Nutrition Survey** see <https://www.cpc.unc.edu/projects/china>

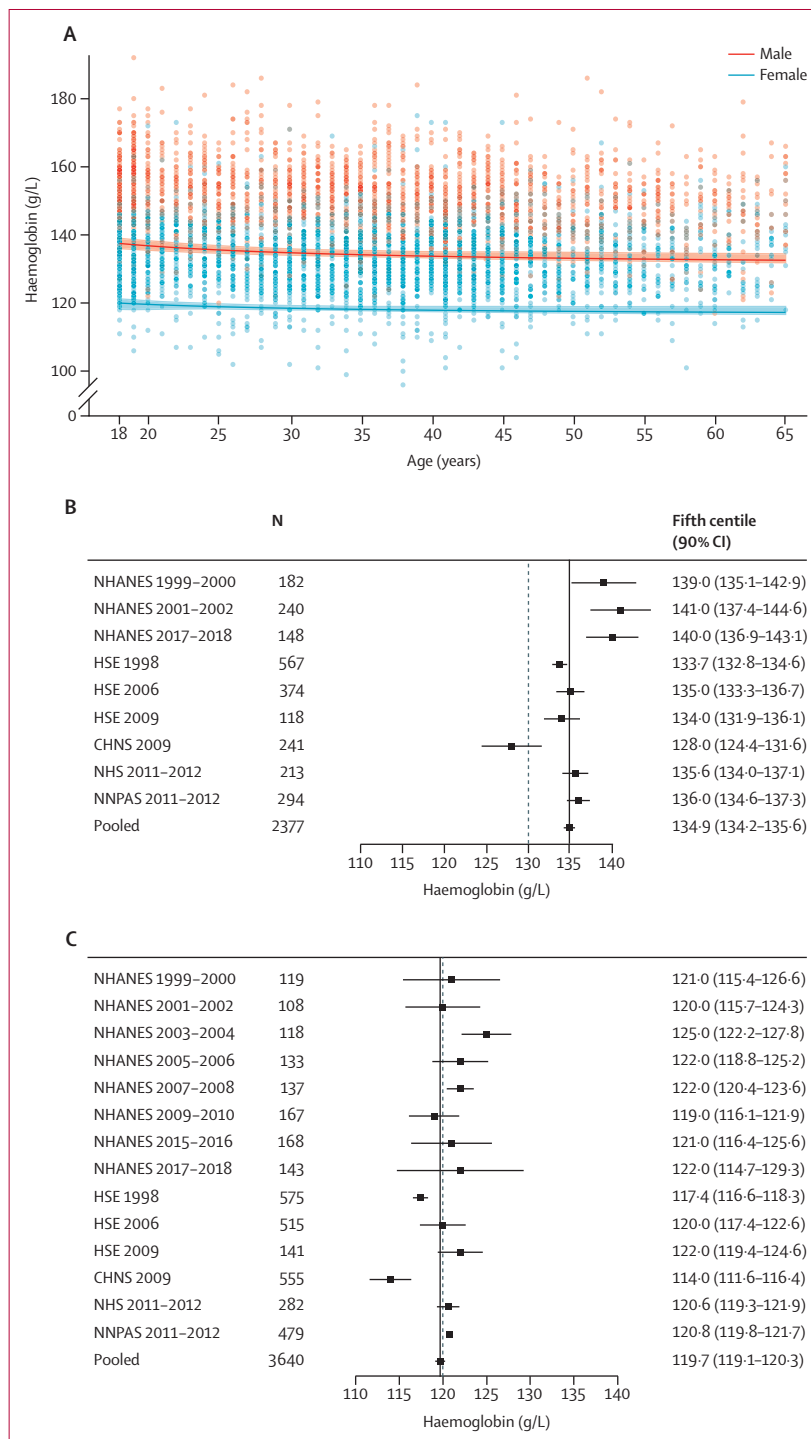


Figure 1: Derivation of haemoglobin thresholds in males and non-pregnant females aged 18–65 years

(A) Continuous thresholds in adult males aged 18–65 years and non-pregnant females aged 18–65 years pooled across the following datasets: NHANES, HSE, and CHNS. The pooled continuous centiles and 90% CIs across data sources were estimated without accounting for the complex survey design and weighting.

(B) Individual study and pooled (fixed effect, black line) estimates of the fifth centile of the haemoglobin distribution in the healthy reference population in male adults aged 18–65 years. (C) Individual study and pooled (fixed effect, black line) estimates of the fifth centile of the haemoglobin distribution in the healthy reference population in non-pregnant female adults aged 18–65 years. The dashed line in panels B and C represents the cutoff for anaemia according to the WHO guideline.⁴ CHNS=China Health and Nutrition Survey. HSE=Health

Survey for England. NHANES=National Health and Nutrition Examination Survey (USA). NHS=National Health Survey (Australia). NNPAS=National Nutrition and Physical Activity Survey (Australia).

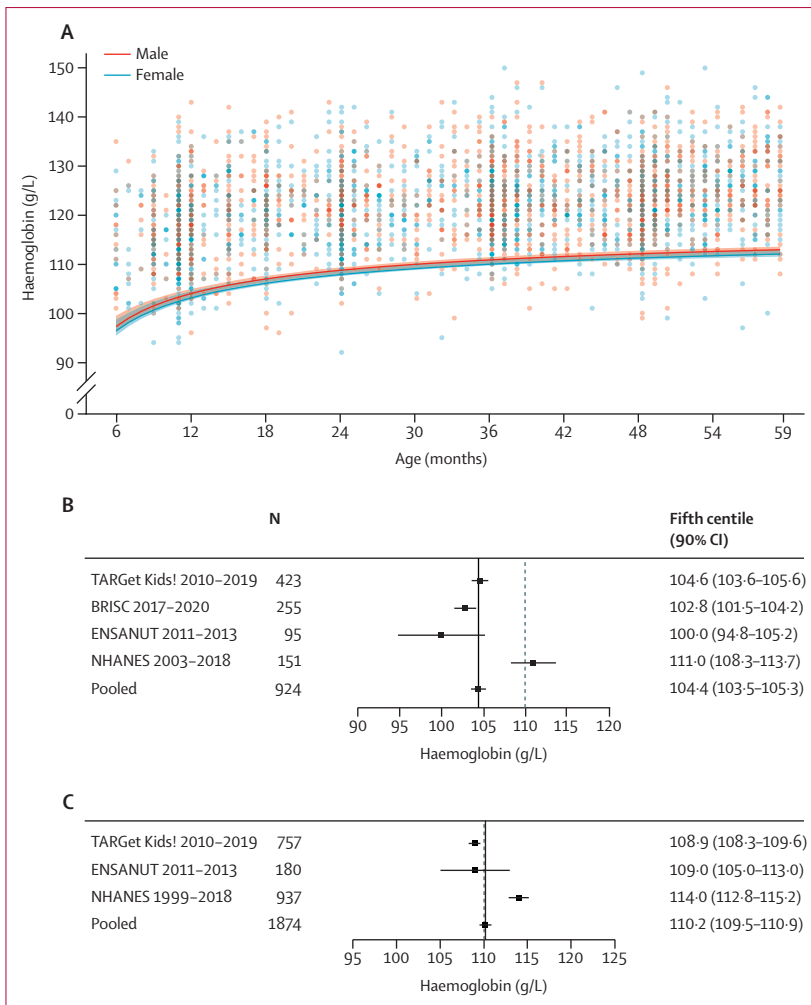


Figure 2: Derivation of haemoglobin thresholds in children aged 6–59 months (A) Continuous thresholds in male and female children aged 6–59 months pooled across the following datasets: TARGet Kids!, BRISC, ENSANUT, and NHANES. The pooled continuous centiles and 90% CIs across data sources were estimated without accounting for the complex survey design and weighting. (B) Individual study and pooled (fixed effect, black line) estimates of the fifth centile of the haemoglobin distribution in the healthy reference population in children aged 6–23 months. (C) Individual study and pooled (fixed effect, black line) estimates of the fifth centile of the haemoglobin distribution in the healthy reference population in children aged 24–59 months. The dashed line in panels B and C represents the cutoff for anaemia according to the WHO guideline.⁴ BRISC=Benefits and Risks of Iron interventions in Children (Bangladesh). ENSANUT=Encuesta Nacional de Salud y Nutrición (Ecuador). NHANES=National Health and Nutrition Examination Survey (USA). TARGet Kids!=Applied Research Group for Kids (Canada).

than 6 months, and our exclusion criteria approach precluded adequate populations of individuals aged 65 years or older from the analyses. Reference populations were derived from multi-ethnic populations in the USA (including individuals identifying as Black or Asian), England (including individuals identifying as Asian), Australia, Canada, and the Netherlands, and populations from China, Bangladesh, and Ecuador. For each survey, demographic characteristics (including age, self-reported ancestry, haemoglobin, iron status, and inflammation) for the overall and reference (healthy) sample are shown in the appendix (pp 26–76). Reasons for exclusion from the reference sample and exclusion

details are provided in the appendix (pp 77–165). Among study participants with non-missing haemoglobin, ferritin, and CRP data, 16 867 (87.6%) of 19 249 adult men, 26 193 (87.8%) of 29 846 adult women, 3429 (78.8%) of 4353 children aged 6–23 months, 3995 (68.1%) of 5869 children aged 24–59 months, 2571 (58.3%) of 4413 children aged 5–11 years, 1517 (57.9%) of 2620 male children and adolescents aged 12–17 years, 3543 (67.0%) of 5288 female children and adolescents aged 12–17 years, 4279 (84.7%) of 5054 women in their first trimester of pregnancy, and 2083 (94.9%) of 2194 women in their second trimester of pregnancy were excluded from the healthy reference sample. Pooled analyses for the fifth centile for haemoglobin concentration in healthy individuals are presented in figures 1–4, while results for the 2.5th centile are available in the appendix (pp 166–68).

Continuous thresholds in males and females based on all available datasets are summarised in the appendix (p 187), indicating periods of change and divergence between sexes in thresholds across the life course from age 6 months to 65 years.

Figure 1 presents estimates of haemoglobin thresholds to define anaemia in adult men (n=2377) and women (n=3640). The data were derived from largely multi-ethnic populations (comprising individuals who self-identified as White, Black, and Asian) from the USA, England, Australia, and China (appendix pp 26–48). Reasons for exclusion of individuals from each dataset are summarised and detailed in the appendix (pp 77–120). Continuous analyses of the data indicate that in adults aged 18–65 years, the fifth centile haemoglobin concentration of the reference population is higher in males than in non-pregnant females (figure 1A; appendix p 187). Pooled analyses indicate that the fifth centile for haemoglobin concentrations in healthy adult men is 134.9 g/L (90% CI 134.2–135.6; figure 1B). In adult non-pregnant women, the fifth centile is 119.7 g/L (119.1–120.3; figure 1C). Sensitivity analyses were done where the ferritin threshold to define iron deficiency was raised from 15 µg/L to 30 µg/L, 45 µg/L, and 100 µg/L (appendix pp 169–70). Haemoglobin values in adult men are 135.7 g/L (90% CI 135.0–136.4) for the 30 µg/L ferritin threshold, 136.0 g/L (135.2–136.8) for the 45 µg/L ferritin threshold, and 137.9 g/L (136.9–138.8) for the 100 µg/L ferritin threshold; and in adult non-pregnant women the haemoglobin values are 121.1 g/L (120.5–121.8) for the 30 µg/L ferritin threshold, 122.2 g/L (121.4–123.0) for the 45 µg/L ferritin threshold, and 117.8 g/L (112.5–123.0) for the 100 µg/L ferritin threshold. In an additional analysis of adult non-pregnant women restricted to those studies that measured haemoglobin in both adult women and men, the fifth centile is 119.2 g/L (90% CI 118.6–119.8).

Figure 2 summarises estimates of haemoglobin thresholds in children aged 6–23 months (n=924) and those aged 24–59 months (n=1874). Data were derived from multi-ethnic populations in Canada,⁴⁹ the USA,

Ecuador,⁴⁸ and Bangladesh (appendix pp 49–55). Reasons for exclusion are summarised in the appendix (pp 121–33). Continuous analysis indicates that thresholds are similar in males and females and hence sex-specific thresholds are not necessary in this age group (figure 2A; appendix p 187). Continuous analyses also indicate an increase in the threshold over the first 5 years of life but especially over the first 2 years (figure 2A; appendix p 187). Discrete thresholds were therefore ascertained in children aged 6–23 months and those aged 24–59 months. Pooled analysis indicates the fifth centile for haemoglobin concentrations in children aged 6–23 months is 104.4 g/L (90% CI 103.5–105.3; figure 2B), and in children aged 24–59 months it is 110.2 g/L (109.5–110.9; figure 2C). Sensitivity analyses were performed in which the ferritin threshold to define iron deficiency was raised, the CRP threshold to define inflammation was lowered from 5 mg/L to 1 mg/L, and in which only children with a mean cell volume higher than 73 fL (aged 6–23 months) or higher than 75 fL (aged 24–59 months) were included. Haemoglobin values for the increased ferritin thresholds were 104.3 g/L (90% CI 103.3–105.2) for a ferritin threshold of 15 µg/L, 103.8 g/L (102.5–105.1) for a ferritin threshold of 30 µg/L, and 103.2 g/L (101.1–105.4) for a ferritin threshold of 45 µg/L in children aged 6–23 months and 110.3 g/L (109.6–111.0) for a ferritin threshold of 15 µg/L, 110.3 g/L (109.2–111.4) for a ferritin threshold of 30 µg/L, and 109.1 g/L (106.8–111.4) for a ferritin threshold of 45 µg/L in children aged 24–59 months; the sample size of children with a ferritin threshold of 100 µg/L was too small. These sensitivity analyses indicate the fifth centile estimate is robust to these parameters, as well as to changes in the statistical methods (appendix pp 171–76).

Figure 3 summarises estimates of haemoglobin thresholds in children aged 5–11 years (n=1839). Data were derived from multi-ethnic populations across Canada, the USA, and China (appendix pp 56–58). Exclusions of individuals from each data set are summarised and detailed in the appendix (pp 134–42). Continuous analyses indicate that thresholds are similar in males and females and hence sex-specific thresholds are not necessary in this age group (figure 3A; appendix p 187). Pooled analyses indicated that the fifth centile for haemoglobin concentrations in this group is 114.4 g/L (90% CI 113.6–115.2; figure 3B). Sensitivity analyses indicated that these thresholds were robust to changes in assumptions and the statistical methods (appendix pp 177–79). Haemoglobin values for the changed thresholds for iron deficiency were 115.7 g/L (90% CI 114.7–116.6) for a ferritin threshold of 30 µg/L and 114.4 g/L (112.6–116.1) for a ferritin threshold of 45 µg/L; the sample size of children with a ferritin threshold of 100 µg/L was too small for an adjusted threshold.

Figure 4 summarises estimates of haemoglobin thresholds in adolescents aged 12–17 years (males: n=1103,

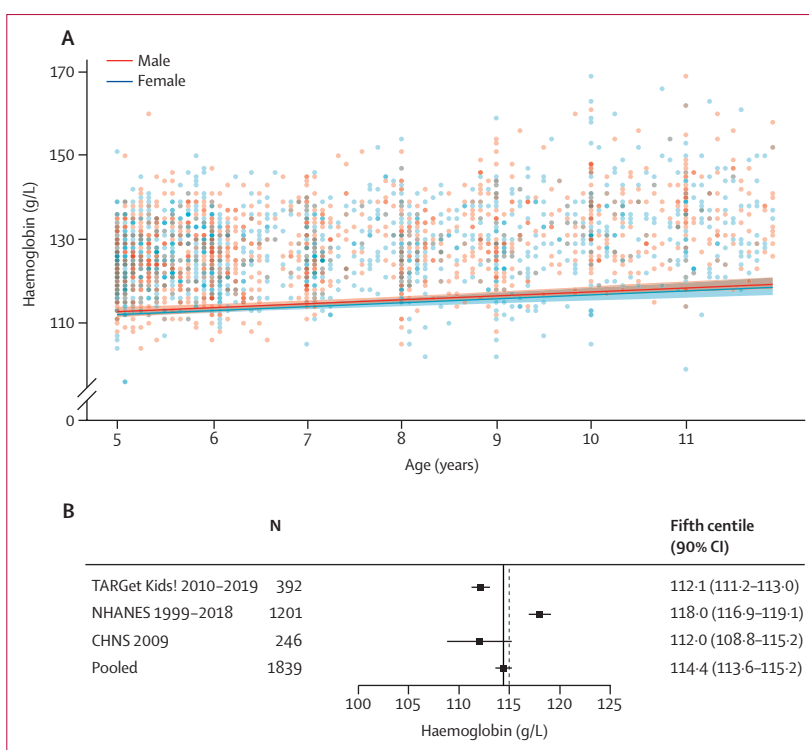


Figure 3: Derivation of haemoglobin thresholds in children aged 5–11 years

(A) Continuous thresholds in male and female children aged 5–11 years pooled across the following datasets: TARGet Kids!, NHANES, and CHNS. The pooled continuous centiles and 90% CIs across data sources were estimated without accounting for the complex survey design and weighting. (B) Individual study and pooled (fixed effect, black line) estimates of the fifth centile of the haemoglobin distribution in the healthy reference population in children aged 5–11 years. The dashed line in panel B represents the cutoff of anaemia according to the WHO guideline.* CHNS=China Health and Nutrition Survey. NHANES=National Health and Nutrition Examination Survey (USA). TARGet Kids!=Applied Research Group for Kids (Canada).

females: n=1741). Data were derived from multi-ethnic populations across the USA, Australia, and China (appendix pp 59–73). Reasons for exclusion of individuals from each dataset are summarised and detailed in the appendix (pp 143–62). Continuous analyses show that thresholds diverge between males and females over this period (figure 4A; appendix p 187), indicating that sex-specific thresholds are necessary in this age group. Pooled analyses indicated that the fifth centile for haemoglobin concentrations in this age group is 128.2 g/L (90% CI 126.4–130.0; figure 4B) in males and 122.2 g/L (121.3–123.1; figure 4C) in females. Sensitivity analyses indicated that these thresholds were robust to changes in assumptions and the statistical methods (appendix pp 180–85). Haemoglobin values for the changed thresholds for iron deficiency were 130.0 g/L (90% CI 127.8–132.3) for a ferritin threshold of 30 µg/L and 131.0 g/L (128.3–133.7) for a ferritin threshold of 45 µg/L in males, and 124.1 g/L (122.8–125.5) for a ferritin threshold of 30 µg/L and 124.0 g/L (122.2–125.8) for a ferritin threshold of 45 µg/L in females; the sample size of children with a ferritin threshold of 100 µg/L was too small. In a post-hoc analysis of female adolescents aged 12–17 years restricted to those studies that measured

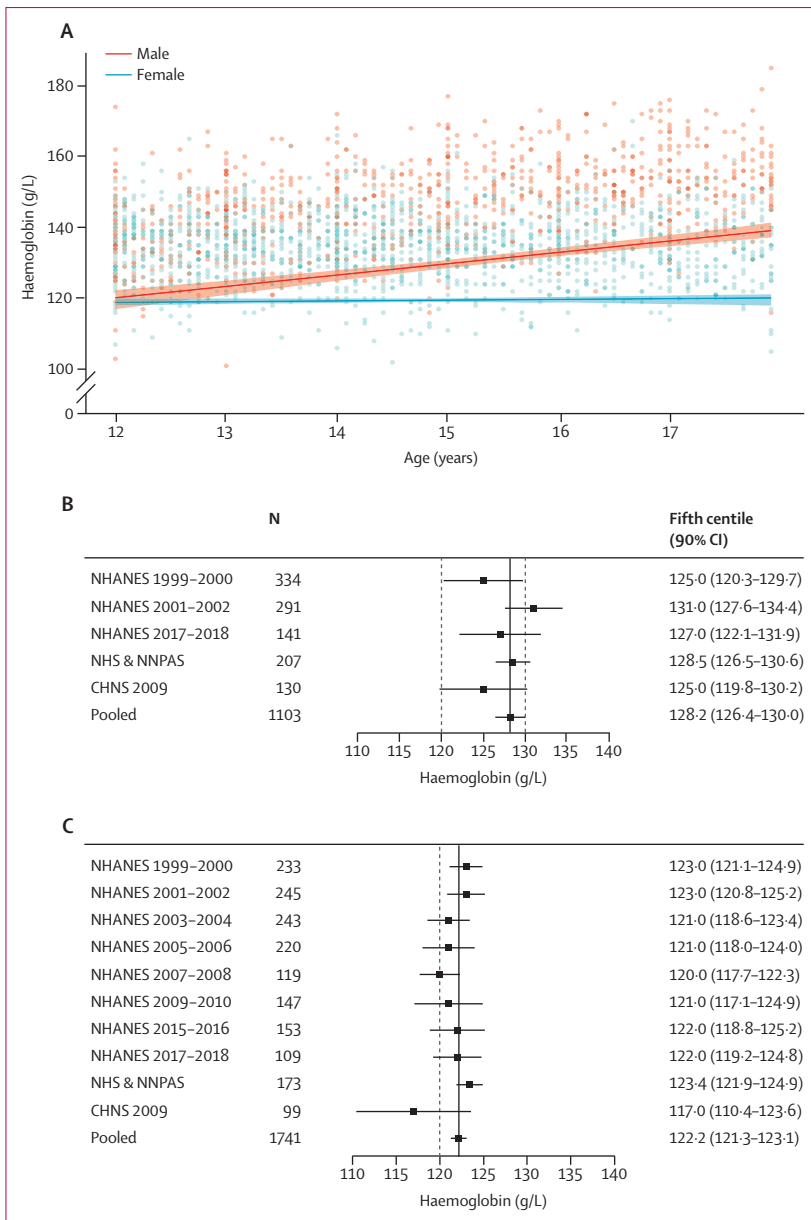


Figure 4: Derivation of haemoglobin thresholds in adolescents aged 12–17 years
 (A) Continuous thresholds in male and female adolescents aged 12–17 years pooled across the following datasets: NHANES and CHNS. The pooled continuous centiles and 90% CIs across data sources were estimated without accounting for the complex survey design and weighting. (B) Individual study and pooled (fixed effect, black line) estimates of the fifth centile of the haemoglobin distribution in the healthy reference population in male adolescents aged 12–17 years. (C) Individual study and pooled (fixed effect, black line) estimates of the fifth centile of the haemoglobin distribution in the healthy reference population in female adolescents aged 12–17 years. The dashed line in panels B and C represents the cutoff for anaemia according to the WHO guideline;⁴ there is a different cutoff for male adolescents aged 12–14 years (120 g/L) and those aged 15 years and older (130 g/L). CHNS=China Health and Nutrition Survey. NHANES=National Health and Nutrition Examination Survey (USA). NHS=National Health Survey (Australia). NNPAS=National Nutrition and Physical Activity Survey (Australia).

haemoglobin in both males and females, the fifth centile is 122.9 g/L (90% CI 121.8–124.1).

Data from Generation R⁵⁰ indicated that the fifth centile for haemoglobin thresholds in healthy pregnant women was 110.3 g/L (90% CI 109.5–111.0) in the first trimester

(n=772), and 105.9 g/L (104.0–107.7) in the second trimester (n=111), with insufficient data for analysis in the third trimester. Participant characteristics are summarised in the appendix (pp 74–76). Exclusions from this dataset are shown in the appendix (pp 163–65). Thresholds were robust to changes in assumptions around the definition of iron deficiency (appendix p 186). Haemoglobin values for the increased thresholds for iron deficiency were 109.9 g/L (90% CI 109.0–110.7) for a ferritin threshold of 30 µg/L and 109.8 g/L (108.8–110.8) for a ferritin threshold of 45 µg/L in the first trimester.

We investigated whether common genetic variation could explain the differences in distribution of haemoglobin concentrations among different ancestry groups (appendix pp 188–94).^{52–56} We found no common SNPs (minor allele frequency [MAF] >0.05) showing large effects on haemoglobin, except for rs13331259 close to the *HBA1* gene and present at a MAF of 0.11 in the African population (appendix p 191). Similarly, a structural variant⁵⁴ in this region (chr16:172001–177200_DEL, *HBA1*, *HBA2/2*) occurred more frequently in individuals of African ancestry (MAF 0.176, appendix p 192), suggesting there might be multiple haemoglobin-associated variants in this region that occur at higher frequencies in African populations, potentially due to positive selection.

We also evaluated rare variation across ancestries in five clinically relevant rare anaemia genes (*HBB*, *KLF1*, *PIEZO1*, *SLC4A1*, and *TMPRSS6*),⁵⁷ and found predicted loss-of-function (pLoF) variants occurred at varying frequencies (appendix pp 193–94). *HBB* pLoF variants occurred more frequently in individuals of east Asian, south Asian, and Middle Eastern ancestries; however, cumulatively the frequency of pLoF variants is very low, less than 0.006. Rare pLoF variants in *TMPRSS6* have the highest prevalence in African and African American populations, with a cumulative pLoF variant frequency of around 0.03.

Discussion

We estimated the fifth centile of haemoglobin concentration in derived reference populations of healthy children, adolescents, and adults, to estimate statistical thresholds to define anaemia across the lifecycle. Analyses of genetic associations with haemoglobin concentration across different ancestries show no evidence of the presence of any non-clinically significant high-frequency variants. Collectively, our analyses guide a single set of global definitions for anaemia.

Statistical definitions for anaemia differ between adult women and men, with the haemoglobin threshold in women robust to sensitivity analyses based on higher ferritin thresholds (eg, <45 µg/L¹²) to indicate iron deficiency. Thresholds between sexes diverge in adolescence, increasing in males. Testosterone concentrations are higher in men than in women, and appear to drive correspondingly higher haemoglobin concentrations.⁵⁸ A direct effect of testosterone has been observed

in previous studies: for example, androgen deprivation in men causes a fall in haemoglobin concentrations that recovers when inhibition is reversed.⁵⁹ Despite the same pre-transfusion haemoglobin concentration, male patients with transfusion-dependent beta-thalassaemia have higher erythropoiesis (eg, soluble transferrin receptor and reticulocyte concentrations) than female patients.⁶⁰ Thus, our data do not support proposals that statistical thresholds to define anaemia should be similar between adult men and women,²⁸ although it is important to recognise that in some studies, similarly reduced haemoglobin concentrations in both sexes have been associated with adverse clinical outcomes.²⁹ Future prospective studies with careful consideration of iron status (beyond ferritin), hormone concentrations, and menstrual blood loss are essential to understand the separation of haemoglobin thresholds by sex.

Our analyses were able to incorporate data from children across a broad age range. Our results highlight the increase in haemoglobin concentration and hence threshold for anaemia that occurs across childhood, and especially in early life. Thresholds based on arbitrary age bands are possibly crude and necessitate jarring changes in the definition of anaemia as children grow and develop. Our data suggest the existing single categorisation at age 6–59 months used by WHO is too broad and suggest at least dividing this population into two smaller subcategories (eg, age 6–23 months and 24–59 months). Definitions of anaemia in male and female children can be combined until adolescence. Recent analyses sought to evaluate the effects of reduced haemoglobin concentrations on pregnancy outcomes, and interestingly indicated an increased risk of preterm birth at haemoglobin thresholds below the statistical thresholds proposed here.⁶¹

Our analysis builds on previous studies using chiefly US-based data to estimate fifth centiles of haemoglobin in iron-replete, non-inflamed populations⁵ by using multinational datasets with the addition of clinical criteria for health, and covers a broad age range. Our findings were robust to variations in criteria for iron deficiency and inflammation, to statistical methods, and to secular health trends. Analyses of population surveys of biochemically iron-replete and non-inflamed individuals across several low-income and middle-income countries have reported substantial heterogeneity of the fifth centile of haemoglobin thresholds between countries in women and preschool children (ie, those aged <5 years).^{36,62,63} These heterogeneous thresholds could relate to high burdens of undetected inflammatory disease in these studies. For example, asymptomatic sub-microscopic *Plasmodium* parasitaemia is highly prevalent in children in sub-Saharan Africa,⁶⁴ while respiratory and diarrhoeal infections are common in infants in south Asia.⁵¹ The high burden of inflammation and unreported frequent clinical illnesses mean that in such settings, a prospectively recruited study with carefully curated inclusion criteria is needed to define a healthy haemoglobin threshold.

Using GWAS we found evidence for the heterogeneity of both common and rare genetic effects on haemoglobin across different ancestries, including two variants proximal to the alpha-thalassaemia genes (*HBA1/HBA2*), which show larger effect sizes than might be expected given their frequency within individuals of African ancestry. These variants are clinically important and must be detectable by a diagnosis of anaemia. We also examined rare variants in genes that are recognised to be clinically relevant causes of inherited anaemia and found these to vary in frequency across populations. Although there was some evidence there might be a higher burden of rare variants with large effects in some ancestral populations, clinically relevant rare variants cumulatively occur at low frequencies and are thus unlikely to influence the overall population-level haemoglobin distribution. The GWAS used in our analyses have high power and incorporate replication. Thus our results do not support the adoption of lower haemoglobin thresholds in populations or individuals of different ancestries, as per recent proposals.⁶⁵ A further limitation of ancestry-specific thresholds is that clinical classification of individuals into ancestral groups is imperfect, as there is substantial genetic diversity within so-called super-populations, and such approaches also do not account for admixed individuals.

Our study has several limitations. First, we defined statistical thresholds on the basis of the fifth centile of the distribution of haemoglobin values in a healthy reference population, rather than functional thresholds that indicate the risk of symptoms or disease; future studies could endeavour to develop functional thresholds. Studies that measured other iron indices, such as transferrin saturation, reticulocyte haemoglobin, soluble transferrin receptor, and hepcidin in clinically healthy populations, might have provided a more holistic picture of iron status.⁸ Scarce data were available on heavy menstrual bleeding, although this common syndrome would be likely to influence haemoglobin through its effects on ferritin concentration, which we measured in every case. Vitamin B12 and folate measurements were not consistently measured in these datasets. Because we sought to use datasets in which clinical data were available and applied conservative exclusion criteria, this approach precluded the use of cross-sectional datasets from many low-income settings where clinical information was scarce and undetected subclinical or recurrent clinical illnesses causing inflammation could have suppressed haemoglobin concentrations, which could lead to under-estimation of thresholds. We were able to use datasets from North and South America, Europe, Australia, south and east Asia, but not sub-Saharan Africa. Most participants were excluded from the healthy reference population because of clinical or biological factors that could affect their haemoglobin concentrations. Importantly, the reference samples generally comprised multi-ethnic populations (including African ancestries). Likewise, as is a common constraint with genetic analyses,

the largest datasets were representative of European ancestry populations, although we specifically included studies recruiting substantial numbers of individuals of south and east Asian and African ancestries. A further limitation of the genetic summaries we have used is that they do not capture the full extent of genetic variation, such as variants on non-autosomal chromosomes. Ideally, a prospective multicentre study could clinically exclude individuals at risk of reduced haemoglobin concentration to define the reference group.^{66,67}

Analysis of thresholds in children younger than 6 months, in pregnant women specifically during the third trimester, and among individuals older than 65 years was not possible due to the scarcity of datasets containing sufficient clinical data and simultaneous collection of haemoglobin and inflammatory and iron parameters. Furthermore, among older populations, datasets contained insufficient numbers of individuals meeting our stringent health criteria. Further work, potentially involving prospective studies with carefully defined clinical inclusion criteria, is needed to more precisely estimate thresholds in these important populations.

In conclusion, this study provides an evidence base to harmonise haemoglobin thresholds to define anaemia in children and adults worldwide and to inform global guidelines.

Contributors

SB, KF, GMM, ED, and SP conceptualised the study, undertook the population statistical analysis together with SZ and JX, and drafted the manuscript. JH, VJ, and MB conceptualised and undertook the genetic study. The Generation R dataset was verified and accessed by SMB, VWV, and RG. The Target Kids! dataset was verified by PP, CB, CK-S, CSB, and JM and accessed by CK-S, SB, and SZ. The National Health and Nutrition Examination Survey (NHANES) and Health Survey for England datasets were accessed by SB, JX, and SZ. The China Health and Nutrition Survey (CHNS) and Encuesta Nacional de Salud y Nutrición (ENSANUT) datasets were accessed by SB and SZ. The Benefits and Risks of Iron interventions in Children (BRISC) dataset was verified by SB and SP, and accessed by SB and SZ. The datasets from the Australian Bureau of Statistics (National Nutrition and Physical Activity Survey and National Health Survey) were accessed by SB, KF, and SZ. The Genes & Health Research Team verified the data of the Genes & Health genetics study and the genetic data for this study were accessed by JH, VJ, and MB. SB, KF, ED, and SP were responsible for the decision to submit the manuscript for publication. SP is the guarantor of this work. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Declaration of interests

SP reports a role as consultant for CSL Vifor and roles as head of the WHO Collaborating Centre for Anaemia Detection and Control, and as co-chair of the Guideline Development Group Meetings for WHO haemoglobin thresholds to define anaemia in clinical medicine and public health. All other authors declare no competing interests.

Data sharing

Study data used for this analysis are either publicly available or were accessed with agreement from primary data collection sources, and will not be available from the study authors.

Acknowledgments

This analysis was funded by WHO and the Bill & Melinda Gates Foundation (INV-050676). SP is funded by the Australian National Health and Medical Research Council (NHMRC; GNT2009047). MB is funded by the NHMRC (GNT1195236). This work was also supported by the Victoria State Government's Operational Infrastructure Support Program and the

NHMRC Independent Research Institute Infrastructure Support Scheme (IRIIS). Funding to support TARGeT Kids! was provided by multiple sources including the Canadian Institutes for Health Research (CIHR), the Hospital for Sick Children Foundation, Toronto, ON, Canada, and the St Michael's Hospital Foundation, Toronto, ON, Canada. The funders had no role in the planning of the study, analysis of data, or the decision to publish this work. Genes & Health has recently been core-funded by Wellcome (WT102627 and WT210561), the UK Medical Research Council (M009017, MR/X009777/1, and MR/X009920/1), Higher Education Funding Council for England Catalyst, Barts Charity (845/1796), Health Data Research UK (for the London substantive site), and research delivery support from the NHS National Institute for Health Research Clinical Research Network (North Thames). Genes & Health has recently been funded by Alnylam Pharmaceuticals and Genomics; and a Life Sciences Industry Consortium of AstraZeneca, Bristol-Myers Squibb, GlaxoSmithKline, Maze Therapeutics, Merck Sharp & Dohme, Novo Nordisk, Pfizer, and Takeda Development Centre Americas. This research uses data from China Health and Nutrition Survey (CHNS). We are grateful for research grant funding from the US National Institutes of Health (NIH), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD; R01 HD30880 and R01 HD38700), the US National Institute on Aging (NIA; R01 AG065357), the US National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; R01 DK104371 and P30 DK056350), the US National Heart, Lung, and Blood Institute (NHLBI; R01 HL108427), an NIH Fogarty Grant (D43 TW009077), the Carolina Population Center (P2C HD050924 and P30 AG066615 since 1989); and support from the China-Japan Friendship Hospital, Ministry of Health (CHNS 2009), the Chinese National Human Genome Center at Shanghai (since 2009), and Beijing Municipal Center for Disease Prevention and Control (since 2011). We thank the many people who have participated in or worked on the studies contributing data to this analysis. We thank Vanessa Pac Soo and Joanna Ling (Methods and Implementation Support for Clinical and Health research Hub, Faculty of Medicine, Dentistry and Health Sciences, The University of Melbourne, Melbourne, Australia) for their assistance with data management. We thank Naomi Von Dinklage (Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia) for administrative support. We thank all members of the Genes & Health Research Team (the full list of members is provided in the appendix p 195). We thank Social Action for Health, Centre of The Cell, members of our Community Advisory Group, and staff who have recruited and collected data from volunteers. We thank the NIHR National Biosample Centre (UK Biocentre), the Social Genetic & Developmental Psychiatry Centre (King's College London), Wellcome Sanger Institute, and the Broad Institute for sample processing, genotyping, sequencing, and variant annotation. We thank Barts Health NHS Trust, NHS Clinical Commissioning Groups (City and Hackney, Waltham Forest, Tower Hamlets, Newham, Redbridge, Havering, and Barking and Dagenham), East London NHS Foundation Trust, Bradford Teaching Hospitals NHS Foundation Trust, Public Health England (especially David Wyllie), Discovery Data Service/Endeavour Health Charitable Trust (especially David Stables), Voror Health Technologies (especially Sophie Don), and NHS England (previously NHS Digital) for GDPR-compliant data sharing backed by individual written informed consent. Most of all, we thank all the volunteers participating in Genes & Health. We thank the National Institute for Nutrition and Health, China Center for Disease Control and Prevention, Beijing Municipal Center for Disease Control and Prevention, and the Chinese National Human Genome Center at Shanghai. We thank all participating children and families for their time and involvement in TARGeT Kids!. We thank the TARGeT Kids! Collaboration for supporting this study. The TARGeT Kids! Collaboration is a primary care practice-based research network and includes practice site physicians, research staff, collaborating investigators, trainees, methodologists, biostatisticians, data management personnel, laboratory management personnel, and advisory committee members.

References

- Weatherall DJ, Hatton C. Anaemia: pathophysiology, classification, and clinical features. In: Warrell DA, Cox TM, Firth JD, eds. *Oxford Textbook of Medicine*, 5th edn. Oxford: Oxford University Press, 2013.
- WHO. Iron deficiency anaemia — report of a study group. Geneva: World Health Organization, 1959.

- 3 WHO. Nutritional anaemias: report of a WHO scientific group [meeting held in Geneva from 13 to 17 March 1967]. 1968. <https://iris.who.int/handle/10665/40707?&locale-attribute=en> (accessed Jan 26, 2024).
- 4 WHO. Nutritional anaemias: tools for effective prevention and control. Geneva: World Health Organization, 2017.
- 5 Beutler E, Waalen J. The definition of anemia: what is the lower limit of normal of the blood hemoglobin concentration? *Blood* 2006; **107**: 1747–50.
- 6 Pasricha SR, Colman K, Centeno-Tablante E, Garcia-Casal MN, Peña-Rosas JP. Revisiting WHO haemoglobin thresholds to define anaemia in clinical medicine and public health. *Lancet Haematol* 2018; **5**: e60–62.
- 7 Colman KS, Wood EM, De La Salle B, Stanworth SJ, Pasricha SR. Heterogeneous hemoglobin lower thresholds in clinical laboratories. *Am J Hematol* 2018; **93**: E142–44.
- 8 Pasricha SR, Tye-Din J, Muckenthaler MU, Swinkels DW. Iron deficiency. *Lancet* 2021; **397**: 233–48.
- 9 Warner MA, Hanson AC, Frank RD, et al. Prevalence of and recovery from anemia following hospitalization for critical illness among adults. *JAMA Netw Open* 2020; **3**: e2017843.
- 10 Shah A, Chester-Jones M, Dutton SJ, et al. Intravenous iron to treat anaemia following critical care: a multicentre feasibility randomised trial. *Br J Anaesth* 2022; **128**: 272–82.
- 11 Sharma AJ, Addo OY, Mei Z, Suchdev PS. Reexamination of hemoglobin adjustments to define anemia: altitude and smoking. *Ann NY Acad Sci* 2019; **1450**: 190–203.
- 12 Ko CW, Siddique SM, Patel A, et al. AGA Clinical Practice Guidelines on the gastrointestinal evaluation of iron deficiency anemia. *Gastroenterology* 2020; **159**: 1085–94.
- 13 Muñoz M, Acheson AG, Auerbach M, et al. International consensus statement on the peri-operative management of anaemia and iron deficiency. *Anaesthesia* 2017; **72**: 233–47.
- 14 Tan YL, Kidson-Gerber G. Antenatal haemoglobinopathy screening in Australia. *Med J Aust* 2016; **204**: 226–30.
- 15 Goldman M, Yi QL, Steed T, O'Brien SF. Changes in minimum hemoglobin and interdonation interval: impact on donor hemoglobin and donation frequency. *Transfusion* 2019; **59**: 1734–41.
- 16 Anand IS, Gupta P. Anemia and iron deficiency in heart failure: current concepts and emerging therapies. *Circulation* 2018; **138**: 80–98.
- 17 Musallam KM, Tamim HM, Richards T, et al. Preoperative anaemia and postoperative outcomes in non-cardiac surgery: a retrospective cohort study. *Lancet* 2011; **378**: 1396–407.
- 18 Caro JJ, Salas M, Ward A, Goss G. Anemia as an independent prognostic factor for survival in patients with cancer: a systemic, quantitative review. *Cancer* 2001; **91**: 2214–21.
- 19 Sullivan PS, Hanson DL, Chu SY, Jones JL, Ward JW, Adult/Adolescent Spectrum of Disease Group. Epidemiology of anemia in human immunodeficiency virus (HIV)-infected persons: results from the multistate adult and adolescent spectrum of HIV disease surveillance project. *Blood* 1998; **91**: 301–08.
- 20 Roubinian NH, Murphy EL, Mark DG, et al. Long-term outcomes among patients discharged from the hospital with moderate anemia: a retrospective cohort study. *Ann Intern Med* 2019; **170**: 81–89.
- 21 Stevens GA, Paciorek CJ, Flores-Urrutia MC, et al. National, regional, and global estimates of anaemia by severity in women and children for 2000–19: a pooled analysis of population-representative data. *Lancet Glob Health* 2022; **10**: e627–39.
- 22 GBD 2021 Anaemia Collaborators. Prevalence, years lived with disability, and trends in anaemia burden by severity and cause, 1990–2021: findings from the Global Burden of Disease Study 2021. *Lancet Haematol* 2023; **10**: e713–34.
- 23 WHO. Global nutrition targets 2025: policy brief series. Dec 20, 2014. <https://www.who.int/publications/i/item/WHO-NMH-NHD-14.2> (accessed Jan 26, 2024).
- 24 UN. SDG indicators - metadata repository. 2023. <https://unstats.un.org/sdgs/metadata/> (accessed Jan 26, 2024).
- 25 WHO. Daily iron supplementation in children 6–23 months of age. Aug 9, 2023. <https://www.who.int/tools/elena/interventions/iron-children-6to23#:~:text=Daily%20iron%20supplementation%20is%20recommended,preventing%20iron%20deficiency%20and%20anaemia> (accessed Jan 26, 2024).
- 26 WHO. WHO guideline: Use of multiple micronutrient powders for point-of-use fortification of foods consumed by infants and young children aged 6–23 months and children aged 2–12 years. Geneva: World Health Organization, 2016.
- 27 Pasricha SR, Gheorghe A, Sakr-Ashour F, et al. Net benefit and cost-effectiveness of universal iron-containing multiple micronutrient powders for young children in 78 countries: a microsimulation study. *Lancet Glob Health* 2020; **8**: e1071–80.
- 28 Weyand AC, McGann PT, Sholzberg M. Sex specific definitions of anaemia contribute to health inequity and sociomedical injustice. *Lancet Haematol* 2022; **9**: e6–8.
- 29 Butcher A, Richards T, Stanworth SJ, Klein AA. Diagnostic criteria for pre-operative anaemia—time to end sex discrimination. *Anaesthesia* 2017; **72**: 811–14.
- 30 Beutler E, West C. Hematologic differences between African-Americans and whites: the roles of iron deficiency and alpha-thalassaemia on hemoglobin levels and mean corpuscular volume. *Blood* 2005; **106**: 740–45.
- 31 Pasupula DK, Reddy PS. When is a South Indian really anemic? *Indian J Clin Biochem* 2014; **29**: 479–84.
- 32 Garcia-Casal MN, Pasricha SR, Sharma AJ, Peña-Rosas JP. Use and interpretation of hemoglobin concentrations for assessing anemia status in individuals and populations: results from a WHO technical meeting. *Ann N Y Acad Sci* 2019; **1450**: 5–14.
- 33 Petitclerc C, Wilding P. International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section. The theory of reference values. Part 2. Selection of individuals for the production of reference values. *J Clin Chem Clin Biochem* 1984; **22**: 203–08.
- 34 WHO. WHO guideline on use of ferritin concentrations to assess iron status in individuals and populations. Geneva: World Health Organization, 2020.
- 35 McKinnon EJ, Rossi E, Beilby JP, Trinder D, Olynyk JK. Factors that affect serum levels of ferritin in Australian adults and implications for follow-up. *Clin Gastroenterol Hepatol* 2014; **12**: 101–08.e4.
- 36 Addo OY, Yu EX, Williams AM, et al. Evaluation of hemoglobin cutoff levels to define anemia among healthy individuals. *JAMA Netw Open* 2021; **4**: e2119123.
- 37 Delgado C, Baweja M, Crews DC, et al. A unifying approach for GFR estimation: recommendations of the NKF-ASN Task Force on reassessing the inclusion of race in diagnosing kidney disease. *Am J Kidney Dis* 2022; **79**: 268–88.e1.
- 38 Clinical and Laboratory Standards Institute. Defining, establishing, and verifying reference intervals in the clinical laboratory; approved guideline—third edition. October 2010. https://clsi.org/media/1421/ep28a3c_sample.pdf (accessed Jan 26, 2024).
- 39 Tukey JW. Exploratory data analysis. Boston, MA: Addison-Wesley, 1977.
- 40 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.
- 41 Cauty AJ, Davison AC. Resampling-based variance estimation for labour force surveys. *Statistician* 1999; **48**: 379–91.
- 42 Woodruff RS. Confidence intervals for medians and other position measures. *J Am Stat Assoc* 1952; **47**: 635–46.
- 43 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.
- 44 Royston P, Wright EM. A method for estimating age-specific reference intervals ('normal ranges') based on fractional polynomials and exponential transformation. *J R Stat Soc Ser A Stat Soc* 1998; **161**: 79–101.
- 45 NHS Digital. Health Survey for England, 2021 part 1. Dec 15, 2022. <https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england> (accessed Jan 26, 2024).
- 46 CDC. National Center for Health Statistics. National Health and Nutrition Examination Survey. 2023. <https://www.cdc.gov/nchs/nhanes/index.htm> (accessed Jan 26, 2024).
- 47 Australian Bureau of Statistics. Australian Health Survey: Users' Guide, 2011–13. <https://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/4363.0.55.001?OpenDocument> (accessed Jan 26, 2024).

- 48 Ramírez-Luzuriaga MJ, Belmont P, Mendieta MJ, et al. Tomo I: Encuesta Nacional de Salud y Nutrición de la población ecuatoriana de cero a 59 años. ENSANUT-ECU 2012. Quito, Ecuador: Ministerio de Salud Pública/Instituto Nacional de Estadísticas y Censos, 2014.
- 49 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.
- 50 Jaddoe VW, Mackenbach JP, Moll HA, et al. The Generation R study: design and cohort profile. *Eur J Epidemiol* 2006; **21**: 475–84.
- 51 Pasricha S-R, Hasan MI, Braat S, et al. Benefits and risks of iron interventions in infants in rural Bangladesh. *N Engl J Med* 2021; **385**: 982–95.
- 52 Chen MH, Raffield LM, Mousas A, et al. Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations. *Cell* 2020; **182**: 1198–213.e14.
- 53 Finer S, Martin HC, Khan A, et al. Cohort profile: East London Genes & Health (ELGH), a community-based population genomics and health study in British Bangladeshi and British Pakistani people. *Int J Epidemiol* 2020; **49**: 20–21i.
- 54 Wheeler MM, Stilp AM, Rao S, et al. Whole genome sequencing identifies structural variants contributing to hematologic traits in the NHLBI TOPMed program. *Nat Commun* 2022; **13**: 7592.
- 55 Hu Y, Stilp AM, McHugh CP, et al. Whole-genome sequencing association analysis of quantitative red blood cell phenotypes: the NHLBI TOPMed program. *Am J Hum Genet* 2021; **108**: 874–93.
- 56 Wang Q, Dhindsa RS, Carss K, et al. Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature* 2021; **597**: 527–32.
- 57 Genomics England. Rare anaemia (version 3.0). March 22, 2023. <https://nhsgms-panelapp.genomicsengland.co.uk/panels/518/v3.0> (accessed Jan 29, 2024).
- 58 Mancera-Soto E, Ramos-Caballero DM, Magalhaes J, Chaves Gomez S, Schmidt WFJ, Cristancho-Mejía E. Quantification of testosterone-dependent erythropoiesis during male puberty. *Exp Physiol* 2021; **106**: 1470–81.
- 59 Weber JP, Walsh PC, Peters CA, Spivak JL. Effect of reversible androgen deprivation on hemoglobin and serum immunoreactive erythropoietin in men. *Am J Hematol* 1991; **36**: 190–94.
- 60 Pasricha SR, Frazer DM, Bowden DK, Anderson GJ. Transfusion suppresses erythropoiesis and increases hepcidin in adult patients with β -thalassemia major: a longitudinal study. *Blood* 2013; **122**: 124–33.
- 61 Ohuma EO, Jabin N, Young MF, et al. Association between maternal haemoglobin concentrations and maternal and neonatal outcomes: the prospective, observational, multinational, INTERBIO-21st fetal study. *Lancet Haematol* 2023; **10**: e756–66.
- 62 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.
- 63 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; **1450**: 105–25.
- 64 Ochwedo KO, Omondi CJ, Magomere EO, et al. Hyper-prevalence of submicroscopic *Plasmodium falciparum* infections in a rural area of western Kenya with declining malaria cases. *Malar J* 2021; **20**: 472.
- 65 Weyand AC, McGann PT. Eliminating race-based reference ranges in haematology: a call to action. *Lancet Haematol* 2021; **8**: e462–66.
- 66 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** (suppl): S15–26.
- 67 Papageorghiou AT, Ohuma EO, Altman DG, et al. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. *Lancet* 2014; **384**: 869–79.