European Journal of Clinical Nutrition (2007) 61, 135–139
 0954-3007/07 $30.00
www.nature.com/ejcn

ORIGINAL ARTICLE

The prevalence of anaemia depends on the definition: an example from the Aboriginal Birth Cohort Study

D Mackerras[1,2] and G Singh[1,2]

[1]*Menzies School of Health Research, Casuarina, Northern Territory, Australia and* [2]*Institute of Advanced Studies, Charles Darwin University, Darwin, Northern Territory, Australia*

Objective: The cutoffs defining anaemia based on haemoglobin and haematocrit recommended by the World Health Organization (WHO) and the United States (US) are different. We compared the prevalences resulting from these definitions and explored the reasons for the variation observed.
Design: A cross-sectional survey within a cohort study of children recruited at birth at Royal Darwin Hospital.
Setting: Subjects were interviewed in their homes or other convenient location and had a blood sample drawn.
Subjects: Five hundred and seventeen Australian Aboriginal children aged 9–13 years inclusive.
Results: Depending on the criterion used, the prevalence of anaemia in the total group ranged between 6% (95% confidence interval: 4.1–8.4%) and 24.4% (20.7–28.3). Using the WHO criteria, girls aged 12–13 years were identified as a target group, having a two- to-six fold higher prevalence than the other groups. When compared to the US criteria, boys aged 12–13 years had the highest prevalence of anaemia, although this was not significant. Simulations show that the WHO cutoffs are based on inconsistent centiles of the age–sex haemoglobin and haematocrit distributions, and that this largely explains the discrepant results.
Conclusions: This variability in definition could lead to inappropriate identification of target groups for intervention programmes.
Sponsorship: National Health and Medical Research Council of Australia, the Colonial Foundation and the Channel 7 Children's Research Foundation of South Australia Inc.
European Journal of Clinical Nutrition (2007) **61,** 135–139. doi:10.1038/sj.ejcn.1602483; published online 19 July 2006

Keywords: anaemia; haemoglobin; haematocrit; surveys

Introduction

Iron-deficiency anaemia is the most common nutritional deficiency (CDC, 1998; WHO, 2001). The definition of anaemia can be based on either haemoglobin or haematocrit concentration. Cutoffs to define anaemia using each parameter are given by both the World Health Organization (WHO) and the US CDC (CDC, 1998; WHO, 2001). There are several important differences between the cutoffs for peripubertal children (Table 1). The WHO criteria groups boys and girls aged 12–13 years together, whereas the US classification separates them. The WHO cutoff for this age lies between the sex-specific American cutoffs. The two sets of cutoffs are derived differently. The current American cutoffs are based on the 5th centile of the distribution from National Health and Nutrition Examination Survey (NHANES) III calculated after excluding those with other evidence of iron deficiency or haemoglobinopathies (CDC, 1998). Because the cutoff for each age–sex group is based on the same point in the distribution, the prevalence in populations with different proportions of ages or boys and girls can be compared directly. By contrast, it is unclear how the current WHO cutoffs were derived, although the value for children 5–14 years refers to a paper from NHANES II that

Correspondence: Dr D Mackerras, Menzies School of Health Research, PO Box 41096, Casuarina, Northern Territory 0811, Australia.
E-mail: Dorothy.Mackerras@foodstandards.gov.au
Guranantor: D Mackerras.
Contributors: DM was the statistical analyst and writer. GS repeated the analysis independently and edited the paper. Both authors participated in data collection.
Received 26 August 2005; revised 13 February 2006; accepted 28 May 2006; published online 19 July 2006

Table 1 Cutoffs defining anaemia based on two haematological parameters from the WHO (WHO, 2001) and the US (CDC 1998) for peri-pubertal children

	WHO cutoffs		*US cutoffs*	
	Haemoglobin	*Haematocrit*	*Haemoglobin*	*Haematocrit*
	<g/dl	*<%*	*<g/dl*	*<%*
Children 9<12 years	<11.5	<34	⩽11.9	⩽35.4
Children 12–13 years	<12	<36	—	—
Boys 12–13 years	—	—	⩽12.5	⩽37.3
Girls 12–13 years	—	—	⩽11.8	⩽35.7

excluded those with iron deficiency (Yip *et al*., 1984) and the haematocrit cutoffs appear to have been calculated from the haemoglobin cutoff using a standard conversion. If the WHO cutoffs for each age are based on different centiles of the distribution, then observed variation in prevalence across age–sex groups could be partly or completely due to definitional artefacts. As the origin of the WHO cutoff is unclear, the expected prevalence is unknown and so interpretation of population data is unclear.

We measured haemoglobin and haematocrit during a survey of peri-pubertal Aboriginal children in Northern Australia. Using these data, we examine the relationships between the current WHO and US cutoffs for this age group and explore the impact of the choice of criterion on the assessment of anaemia prevalence.

Subjects and methods

The Aboriginal Birth Cohort Study is described in detail elsewhere (Sayers *et al*., 2003). Briefly, 686 out of 1238 children born at the Royal Darwin Hospital between January 1987 and March 1990 with mothers self-identifying as Aboriginal were recruited into the study. Non-recruitment was mainly owing to unavailability of the neonatologist assessing gestational age using the Dubowitz method and/or inability to locate the mother to inform her about the study (Sayers *et al*., 2003). There were no significant differences in the mean birth weight, birth weight frequencies or sex ratio between those recruited and not recruited. Recruited children were re-seen between December 1998 and March 2001 and examined for a variety of factors (Mackerras *et al*., 2003; Sayers *et al*., 2003). A venous blood sample was drawn after the application of a topical anaesthetic. A full blood count including haemoglobin and packed cell volume (haematocrit) was carried out using a MAXM Hematology Flow Cytometer (Beckman Coulter, Fullerton, USA) but no biochemical tests of iron deficiency were carried out. Growth was assessed by calculating height-for-age and weight-for-age *z*-scores based on the 1978 references using EpiInfo 2000 (Centers for Disease Control and Prevention, GA, USA). The study was approved by the Joint Institutional Ethics Committee of the Menzies School of Health Research and the Royal Darwin Hospital. This Committee has an Aboriginal Sub-Committee with veto power. Parents or guardians gave written informed consent and children gave verbal consent.

Statistical methods

Anaemia was defined in four ways, using the WHO and US cutoffs based on each of haemoglobin and haematocrit (CDC, 1998; WHO, 2001). Differences in the prevalence of anaemia among the age–sex groups were tested using χ^2 and an alpha of 0.05.

We calculated the *z*-score for each of the cutoffs using the appropriate age- and sex-specific distribution from the NHANES II (Yip *et al*., 1984), which was referenced in the document (WHO, 2001) to investigate whether the observed variation in prevalence was owing to variation in the location of the cutoffs in the underlying frequency distribution. In that paper (Yip *et al*., 1984), the medians, 2.5th and 97.5th centiles are given. As the ranges were not exactly even around their respective medians, we assumed that each distribution could be regarded as composed of half of two Normal distributions, each with a slightly different standard deviation. We estimated the lower standard deviation of the various age–sex distributions by (median−2.5th centile value)/1.96. Then, we calculated the age–sex-specific *z*-score, and its associated centile, for each of the WHO cutoffs using the appropriate medians and estimated standard deviations.

The current US cutoffs are based on the NHANES III distribution (CDC, 1998). Therefore, *z*-scores for the WHO haemoglobin cutoffs were also calculated using the mean and standard deviations for age–sex-specific haemoglobin distributions from this survey (Looker *et al*., 1997) to remove any effects owing to using different reference populations. The centile associated with each *z*-score was then derived. The equivalent process was not repeated for haematocrit because we were unable to find a description of haematocrit distributions from this survey on the website containing other reports (http://www.cdc.gov/nchs/products/pubs/pubd/series/sr11/ser11.htm). All analyses were carried out using Stata 8 (College Station, TX, USA) and Excel.

Results

Of the 686 children initially recruited, 572 were seen and blood samples were obtained on 523 children. For the current analysis, we excluded four children aged 8 years and two children aged 14 years to reduce any effects relating to small numbers at extreme ages. Demographic, growth and haematological characteristics of the children are shown (Table 2). The mean height and weight was approximately half a standard deviation below the reference median and this is consistent with other reports of poor growth in Aboriginal children in remote areas (Cunningham and

Table 2 Characteristics of the 517 children by age–sex group, Aboriginal Birth Cohort Study, Australia

	N	*Mean*	*s.d.*	*Median*		*IQR*
Children aged 9–11 years						
Sex (% male)	331	48.6	—	—	—	
Haemoglobin (g/dl)	331	12.6	1.0	12.5	12.0	13.1
Haematocrit (%)	331	38.0	2.9	38.0	36.0	39.0
Height-for-age *z*-score	330	−0.5	1.1	−0.6	−1.3	0.1
Weight-for-age *z*-score	330	−0.7	1.3	−1.0	−1.6	0.1
Boys aged 12–13 years						
Haemoglobin (g/dl)	110	13.1	1.0	13.0	12.4	13.6
Haematocrit (%)	110	39.5	3.1	39.5	37.0	41.0
Height-for-age *z*-score	110	−0.5	1.1	−0.8	−1.4	0.0
Weight-for-age *z*-score	110	−0.7	1.2	−1.0	−1.7	0.0
Girls aged 12–13 years						
Haemoglobin (g/dl)	76	12.6	1.1	12.6	11.9	13.3
Haematocrit (%)	76	38.1	3.1	38.0	36.0	40.0
Height-for-age *z*-score	76	−0.4	1.2	−0.4	−1.1	0.4
Weight-for-age *z*-score	76	−0.3	1.2	−0.3	−1.2	0.4

Abbreviation: IQR, inter-quartile range.

Table 3 Prevalence of anaemia using cutoffs defined for different parameters by different authorities, by age–sex group, Aboriginal Birth Cohort Study, Australia

Parameter	*WHO (2001)*		*US (CDC, 1998)*	
	*Prevalence (%)**	*95% CI*	*Prevalence (%)*#	*95% CI*
All children aged 9–13 years				
Haemoglobin	13.2	10.4–16.4	24.4	20.7–28.3
Haematocrit	6.0	4.1–8.4	19.7	16.4–23.4
Children aged 9–11 years				
Haemoglobin	9.4	6.5–13.0	23.3	18.8–28.2
Haematocrit	3.3	1.7–5.9	17.8	13.9–22.4
Boys aged 12–13 years				
Haemoglobin	13.6	7.8–21.5	30.0	21.6–39.5
Haematocrit	4.5	1.5–10.3	25.5	17.6–34.6
Girls aged 12–13 years				
Haemoglobin	28.9	19.1–40.9	21.1	12.5–31.9
Haematocrit	19.7	11.5–30.5	19.7	11.5–30.5

*$P<0.001$ for difference among the three age–sex groups for both parameters.
#$P>0.2$ for difference among the three age–sex groups for both parameters.

Mackerras, 1998). None of the children had macrocytosis on visual inspection of the slide.

The four different cutoffs yielded an apparent prevalence of anaemia ranging from 6.0 to 24.4% overall and from 3.3 to 30.0% within the various subgroups (Table 3, Figure 1). The cutoff based on haemoglobin yields a higher prevalence than the cutoff based on haematocrit within each source, although this was not significant for the US cutoffs for boys and girls aged 12–14 years. However, pattern between the age and sex subgroups differs according to source of the cutoffs. Using the WHO cutoffs, the prevalence varies significantly and is much higher in the 12 to 13-year-old girls than in the other groups. However, there is no significant variation in the prevalence of anaemia among the three groups when the US cutoffs are used.

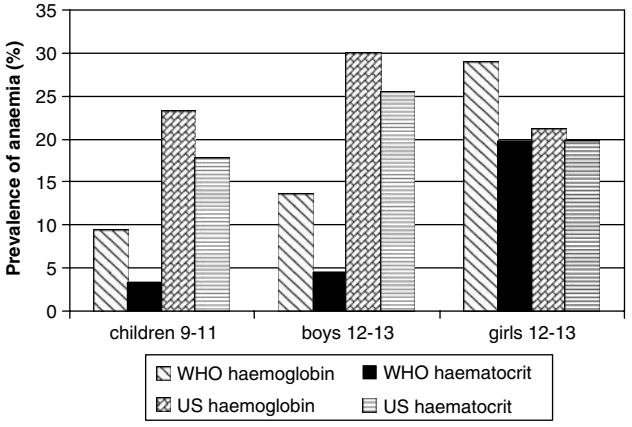

Figure 1 Prevalence of anaemia based on the WHO (2001) and US (CDC, 1998) cutoffs for haemoglobin and haematocrit, by age and sex, Aboriginal Birth Cohort Study, Australia.

The reason for these differences is apparent when the WHO cutoffs are compared to the NHANES II or III distributions (Table 4). The WHO cutoffs do not have a constant *z*-score either within age–sex groups for one index or between the two indices. Calculations (not shown) indicate that this variation also occurs in other age/sex groups. Consequently, the expected prevalence of values below the cutoffs would vary across the groups even if the study population had an identical distribution to the reference distribution.

Discussion

In our population, the prevalence of anaemia depends on the definition that is used. For the total group, it could be presented to policy makers as low (6%) or as high (24.4%). In addition, the choice of definition alters which group is identified as the highest priority for targeting. Using either of the WHO criteria, girls aged 12–13 years would be the target group, whereas all three groups would be targeted if the US criteria are used.

A major reason for the discrepancies is that the WHO cutoffs defining anaemia appear to be set at different centiles for different age–sex groups and for the two parameters. This introduces confounding by age and sex when comparing the prevalence of anaemia within populations and between populations. Expressed as centiles, the WHO cutoffs vary between the 2.5th centile and the 12th centile when compared to the NHANES II distribution calculated after excluding those with indicators of iron deficiency. Consequently, in a population with the same haematocrit distribution as the non-iron-deficient reference, girls have

Table 4 Location of the WHO anaemia cutoffs on the (a) NHANES II and (b) NHANES III distributions

(a)

	NHANES II distribution (Yip et al., 1984)				*(WHO, 2001) cutoffs*		
	Age (years)	*Median*	*2.5th centile*	*Estimated lower s.d.*	*Value*	*z-score*	*Centile*
Haemoglobin							
Children	9–11	13.2	11.4	0.918	<11.5	−1.85	3.2
Boys	12–14	14.0	12.0	1.020	<12.0	−1.96	2.5
Girls	12–14	13.4	11.5	0.969	<12.0	−1.44	7.4
Haematocrit							
Children	9–11	38.4	34	2.245	<34	−1.96	2.5
Boys	12–14	40.5	35	2.806	<36	−1.60	5.4
Girls	12–14	39.0	34	2.551	<36	−1.18	12.0

(b)

	NHANES III haemoglobin distribution (Looker et al., 1997)			*(WHO, 2001) cutoffs*		
	Age (years)	*Mean*	*s.d.*	*Value*	*z-score*	*Centile*
Children	9–11	13.09	0.79	<11.5	−2.01	2.2
Boys	12–14	14.24	1.00	<12.0	−2.24	1.3
Girls	12–14	13.43	0.93	<12.0	−1.54	6.1

Abbreviations: NHANES, National Health and Nutrition Examination Survey; WHO, World Health Organization.

an expected prevalence of 12% compared to an expected prevalence of 2.5% in children aged 9–11 years. However, the discrepancy becomes exaggerated as the study population average departs from the reference average. For example, suppose the haemoglobin distribution in a hypothetical population of boys and girls aged 12–14 years is shifted to the left by 0.5 standard deviations for each sex, but that the standard deviations are the same as those of the NHANES II study shown in Table 4. For boys, the WHO cutoff, which is on the 2.5th centile (z-score $= -1.96$) of the NHANES II reference, will lie at a z-score of -1.46 (the 7.2nd centile) of the hypothetical population distribution, but for girls, the WHO cutoff, which lies at the 7.4th centile (z-score -1.446) of the reference, will lie at a z-score of -0.946 (the 17.2nd centile) of the hypothetical distribution. In other words, the WHO cutoffs will yield a prevalence of 7.2% in boys and 17.2% in girls, even though the haemoglobin distributions have been equally shifted and so the prevalence of low values should be the same. Similarly, if the mean haemoglobins are reduced by 1.0 standard deviation but the standard deviations remain the same as the reference populations, then the WHO cutoffs will yield prevalences of 16.9 and 32.8% in boys and girls, respectively. These simulations are not extreme. In our population, the mean haemoglobin for boys aged 12–13 years was more than 1 standard deviation below the median/mean from both reference surveys, whereas the mean for children aged 9–11 years was approximately half a standard deviation below the reference averages. Thus, the variation in anaemia prevalence based on the WHO cutoffs observed in our population subgroups can be largely attributed to the inconsistent location of the cutoffs, rather than any true variation between the groups.

Using the US cutoffs, the anaemia prevalence differed when based on haemoglobin and haematocrit, even though both cutoffs are based on the 5th centile of the appropriate distribution. This is not surprising as the two parameters are not the same, although they are used interchangeably and have a very strong correlation. Haemoglobin is the oxygen-binding protein present in the red blood cells, whereas haematocrit is the proportion of the blood volume occupied by red blood cells. However, the prevalence pattern in the three groups was the same for both indicators.

The finding that the WHO cutoffs vary in their location on the reference population distribution has some important implications. Firstly, comparisons of different population subgroups may lead to incorrect conclusions about which are the high priority target groups. Secondly, comparisons between populations that span more than one age–sex group will be affected because they will depend on the age–sex distribution within each group. Similar observations regarding measurement and reporting of child growth (Sykes, 1977) lead to the development of charts with criteria based on constant centiles for all ages, albeit these charts are somewhat hampered by their underlying data (Dibley *et al*., 1987). Thirdly, an advantage of using, and specifying the basis of, the criteria is that survey data can be correctly interpreted and targets for programme performance can be set appropriately. If the criterion is based on the 5th centile, like the US cutoffs described in this paper, then the expected prevalence is 5% and only a prevalence substantially higher than 5% indicates that a problem exists. Similarly, it would be irrational to expect a successful intervention to yield a prevalence of less than 5% if the criterion was set at the 5th centile. Therefore, we conclude that our total study popula-

tion has an excess prevalence of anaemia of 19.5% using haemoglobin (i.e. 24.5–5% = 19.5% from Table 3), or 14.7% (i.e. 19.7–5%) using haematocrit, above the 5% expected by definition.

Our findings raise questions about the way that some commonly used cutoffs to define anaemia have been set. These could potentially lead to incorrect exclusion of subgroups from intervention programmes. We were not able to explore other parameters related to iron status in a similar manner, but this should be done. In the meantime, we suggest that caution should be exercised when comparing the prevalence between populations and that, in some populations, all peri-pubertal children, rather than solely girls in early teenage, may need targeting. In the meantime, what should authors do? We would suggest exploring whether using cutoffs defined by different groups to determine whether it alters the interpretation of the findings. For example, when the cutoffs are similar and the prevalence is low and the age band narrow, the interpretation can be the same (Mackerras *et al*., 2004). However, when the interpretation differs, we think results based on cutoffs that are unbiased across age–sex groups, such as the US cutoffs, are the most appropriate interpretation. Local criteria should only be used if they are derived in an unbiased way, but may be important to develop if the local assay method is different from that used to derive other cutoffs. We think that it would be useful for authors to report prevalence of anaemia using both the US and WHO cutoffs to generate discussion about whether it is time to revise the definitions.

Acknowledgements

We would like to acknowledge the children and their mothers and those who have worked on the study over the years. Both authors contributed to data collection and the design, analysis and drafting of the paper. There are no conflicts of interest.

References

Centers for Disease Control and Prevention (1998). Recommendations to prevent and control iron deficiency in the United States. *Morb Mortal Weekly Rep* **47**, www.cdc.gov/mmwr/PDF/rr/rr4703.pdf.

Cunningham J, Mackerras D (1998). *Overweight and Obesity, Indigenous Australians*, Australian Bureau of Statistics Cat. No. 4702.0. Australian Bureau of Statistics: Canberra.

Dibley MJ, Goldsby JB, Staehling NW, Trowbridge FL (1987). Development of normalized curves for the international growth reference: historical and technical considerations. *Am J Clin Nutr* **46**, 736–748.

Looker AC, Dallman PR, Carroll MD, Gunter EW, Johnson CL (1997). Prevalence of iron deficiency in the United States. *JAMA* **27**, 973–976.

Mackerras DEM, Hutton SI, Anderson PR (2004). Haematocrit levels and anaemia in Australian children aged 1–4 years. *Asia Pac J Clin Nutr* **13**, 330–335.

Mackerras DEM, Reid A, Sayers SM, Singh GR, Bucens IK, Flynn KF (2003). Growth and morbidity in children in the Aboriginal Birth Cohort Study: the urban–rural differential. *Med J Aust* **178**, 56–60.

Sayers SM, Mackerras D, Singh G, Bucens I, Flynn K, Reid A (2003). An Australian Aboriginal birth cohort: a unique resource for a life course study of an indigenous population. A study protocol. *BMC Int Health Hum Rights* **3**, 1 (http://www.biomedcentral.com/1472-698X/3/1).

Sykes PA (1977). Errors arising through using the Harvard tables and percentage levels of median weight-for-age in assessing nutritional status. *Arch Dis Child* **52**, 391–394.

United Nations Children's Fund, United Nations University, World Health Organisation (2001). *Iron Deficiency Anaemia Assessment, Prevention and Control. A Guide for Programme Managers*. World Health Organisation: Geneva.

Yip R, Johnson C, Dallman PR (1984). Age-related changes in laboratory values used in the diagnosis of anemia and iron deficiency. *Am J Clin Nutr* **39**, 423–427.