Performance of the Haemoglobin Colour Scale in diagnosing severe and very severe anaemia

Antonio Montresor\(^1\), Mahdi Ramsan\(^2\), Nassor Khalfan\(^2\), Marco Albonico\(^3\), Rebecca J. Stoltzfus\(^4\), James M. Tielsch\(^5\) and Lorenzo Savioli\(^1\)

\(^1\) World Health Organization, Geneva, Switzerland
\(^2\) Pemba Public Health Laboratory Ivo de Carneri, Chake Chake, Pemba, Tanzania
\(^3\) Ivo de Carneri Foundation, Milan, Italy
\(^4\) Division of Nutritional Science, Cornell University, Ithaca, NJ, USA
\(^5\) Center for Human Nutrition, Department of International Health, Bloomberg School of Public Health, Baltimore, MD, USA

Summary

OBJECTIVE To assess the accuracy of Haemoglobin Colour Scale (HCS) in identifying severely anaemic [haemoglobin (Hb) <7 g/dl] and very severely anaemic (Hb <5 g/dl) individuals, and to compare the performances of a group of health workers using HCS after training with a standard method.

METHOD The study consisted of two parts. In part 1, the performance of HCS was compared with clinical examination in a hospital population of which more than 450 individuals were severely anaemic and more than 120 very severely anaemic. Part 2 of the study was conducted in eight dispensaries where the performances of 13 health workers using the colour scale were compared with the performances of eight other health workers using clinical signs to estimate Hb.

RESULTS The colour scale was 92% sensitive for both severe anaemia and very severe anaemia and performed better than clinical examination. Health workers who used the colour scale did better in identifying anaemic and severely anaemic individuals, than those who used clinical examination.

CONCLUSIONS The colour scale improves health workers’ capacity to identify severely anaemic individuals and could be used as a basis for referral. Although the diagnostic accuracy of the workers using HCS varied widely, in most cases they did significantly better than those who used clinical investigation alone.

keywords severe anaemia, haemoglobin colour scale, clinical signs, pallor

Introduction

The accurate measurement of haemoglobin (Hb) allows identifying and treating anaemic individuals, thus preventing the severe consequences of this disorder (Shulman et al. 2001). Unfortunately, in the peripheral health services of least developed countries, where the capacity to identify anaemic individuals is most needed, Hb is normally estimated by clinical examination, which requires long training and vast clinical experience to be accurate (Gjorup et al. 1986; Luby et al. 1995).

The haemoglobin colour scale (HCS) was developed by Stott and Lewis (1995) to improve Hb estimation in settings where laboratory facilities are not available. Its performance was evaluated in several settings (Beales 1997; Münster et al. 1997; Lewis et al. 1998; Van den Broek et al. 1999; Ingram & Lewis 2000) with promising results. In the largest field evaluation \((n = 530)\) conducted by Montresor et al. (2000) the tool gave a sensitivity of 85.2% for anaemia (Hb < 11 g/dl) and of 73.6% for severe anaemia (Hb < 7 g/dl). The measurements were conducted in a population in which very few individuals had severe anaemia \((n = 19)\) and, therefore, the performance of HCS at low levels of Hb was considered to be insufficiently evaluated.

The present study was conducted on Pemba Island, Zanzibar, United Republic of Tanzania, and was composed of two parts: part 1 aimed at assessing the capacity of HCS to identify severely (Hb < 7 mg/dl) and very severely anaemic patients (Hb < 5 mg/dl) and its usefulness as a basis for referral for transfusion and to compare the performance of HCS with the results obtained with the assessment of clinical signs of anaemia (CSA) (pallor of conjunctiva, palms and nail bed) while part 2 was conducted to assess the range of performance of the HCS method when implemented by different primary health care workers trained with a standard protocol.
Materials and methods

The first part of the study was conducted, between November 2000 and April 2001, in three referral health units (Chake-Chake Hospital, Wete Hospital, and Micheweni Cottage Hospital) among new admissions to the paediatric and maternity wards. The second part of the study was conducted, between May and June 2001, in eight dispensaries in Pemba Island evaluating Hb among pregnant women attending the antenatal care services. In both parts of the study, Hb concentration was assessed by three methods: (i) by digital haemoglobinometer; (ii) by the HCS; and (iii) by assessment of CSA.

Capillary blood was collected from a single finger-prick performed with a safety flow lancet (Vacutainer System, Becton Dickinson, NJ, USA). The blood from the finger-prick was collected without compressing the finger. The first drop of blood was discarded, the second was used for HCS analysis and the third for the HemoCue determination. The measurement made by digital haemoglobinometer (HemoCue AB, Angelhom, Sweden) was considered the gold standard. The accuracy of the instrument was checked every day using standard controls.

The HCS (Stott & Lewis 1995) consists of six standard colours varying from pale to dark red, corresponding to haemoglobin levels of 4, 6, 8, 10, 12 and 14 g/dl. To estimate the Hb value, a drop of blood absorbed on a paper test strip (Whatman 31 ETHchr) is matched against the scale standards. Readers can also report intermediate values (e.g. 11 g/dl, 13 g/dl,) if they judge that the colour falls between two standards. Sensitivity and specificity of the HCS was estimated in order to identify anaemic, severely anaemic and very severely anaemic individuals and the difference between the gold standard and the HCS estimation was calculated for each individual. CSA was assessed at three anatomical sites: nail bed, inferior conjunctiva and palm. For each site the observation was coded as ‘normal’, ‘pale’ or ‘very pale’. Estimation of the sensitivity and specificity was carried out, to identify: (1) anaemic individuals as ‘pale’; (2) anaemic individuals as ‘pale’ or ‘very pale’; (3) severely anaemic individuals as ‘very pale’; (4) severely anaemic individuals as ‘pale’ or ‘very pale’; (5) very severely anaemic individuals as ‘very pale’ and (6) very severely anaemic individuals as ‘pale’ or ‘very pale’. This was performed with the aim of evaluating if the possibility of classifying pallor in three categories would have improved the identification of severe or very severe anaemia.

In each health facility, a trained member of the Pemba Public Health Laboratory staff determined the Hb by HemoCue, one or two health workers were recruited to estimate patient Hb with HCS and one health worker was recruited to estimate patient Hb by CSA. These staff received a 2-day standard training on the use of HCS based on the material developed by WHO (included in the HCS package) and on identification of clinical signs of anaemia. To avoid bias in readings, all Hb evaluations were blinded (without examining the patient’s admission chart or knowing the results of the other Hb estimations).

In the second part of the study, the performances of the health workers was classified in three categories: excellent, adequate and poor. For anaemia, because of the low cost of therapy, the screening method should help to identify as many anaemic individuals as possible. Only high sensitivity identification was considered to be acceptable: excellent [sensitivity 80% and positive predictive value (PPV) 80%]; adequate (sensitivity 80% and PPV <80%) and poor (sensitivity <80%).

For severe anaemia, because of the relatively high cost of the intervention, the capacity to identify as many severely anaemic individuals as possible and at the same time reducing the number of the false positives was considered essential for the screening method; high sensitivity and high specificity are therefore considered the leading indicators: excellent (sensitivity ≥80% and specificity ≥80%); adequate (sensitivity 60% to <80% and/or specificity 60% to <80%); poor (sensitivity <60% and/or specificity<60%).

Availability of iron supplements and anthelminthic drugs was ensured for the treatment of anaemic patients participating in both parts of the study. The clinical decision on the need for medical treatment, however, was left to the medical staff in charge in the health units, according to the standard treatment protocol recommended by the Ministry of Health of Zanzibar.

The study received the approval of the ethical committees of the Ministry of Health of Zanzibar, Johns Hopkins Bloomberg School of Public Health and the World Health Organization.

Results

Part 1

The sample was composed of 1633 individuals (990 from paediatric departments, mean age 1.7 years and 643 from the obstetric/gynaecology department, mean age 29.9 years). In 60 cases, one of the readings was incorrectly recorded and therefore discarded during data entry. The mean Hb of the remaining 1573 individuals measured by HemoCue, was 8.45 g/dl (SD = 2.48, range 2.3–16.8 g/dl); 1330 individuals (85%) were classified anaemic, 459 (29%) severely anaemic and 141 (9%) very severely anaemic. More than 55% of the HCS estimates were within 1 g/dl of the gold standard and more than 83% were within 2 g/dl. Sensitivity, specificity, PPV and
negative predictive value (NPV) for anaemia, severe anaemia and very severe anaemia the two methods are shown in Table 1.

### Table 1 Performance in the identification of anaemic individuals (Hb < 11 g/dl), severely anaemic individuals (Hb < 7 g/dl) and very severely anaemic (Hb < 5 g/dl) using Haemoglobin Colour Scale or clinical signs of anaemia

<table>
<thead>
<tr>
<th>Identification of anaemic (Hb &lt; 11 g/dl)</th>
<th>Performances of Haemoglobin Colour Scale</th>
<th>Performances of clinical sign of anaemia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
<td>Absolute numbers</td>
<td>Percentage</td>
</tr>
<tr>
<td></td>
<td>1573</td>
<td>84</td>
</tr>
<tr>
<td>Number of anaemics in the sample</td>
<td>1330</td>
<td>97 (96–97)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>1291/1330</td>
<td>97 (96–97)</td>
</tr>
<tr>
<td>Specificity</td>
<td>100/243</td>
<td>41 (34–47)</td>
</tr>
<tr>
<td>PPV</td>
<td>1291/1434</td>
<td>90 (88–91)</td>
</tr>
<tr>
<td>NPV</td>
<td>100/139</td>
<td>71 (63–79)</td>
</tr>
<tr>
<td>Identification of severe anaemics (Hb &lt; 7 g/dl)</td>
<td>Absolute numbers</td>
<td>Percentage</td>
</tr>
<tr>
<td></td>
<td>459</td>
<td>29</td>
</tr>
<tr>
<td>Number of severe anaemics in the sample</td>
<td>423/459</td>
<td>92 (89–94)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>956/1114</td>
<td>86 (84–88)</td>
</tr>
<tr>
<td>Specificity</td>
<td>423/572</td>
<td>74 (70–77)</td>
</tr>
<tr>
<td>PPV</td>
<td>965/1001</td>
<td>96 (95–97)</td>
</tr>
<tr>
<td>Identification of very severe anaemics (Hb &lt; 5 g/dl)</td>
<td>Absolute numbers</td>
<td>Percentage</td>
</tr>
<tr>
<td></td>
<td>141</td>
<td>8</td>
</tr>
<tr>
<td>Number of very severe anaemics in the sample</td>
<td>130/141</td>
<td>92 (84–94)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>1299/1432</td>
<td>90 (88–91)</td>
</tr>
<tr>
<td>Specificity</td>
<td>130/263</td>
<td>49 (43–55)</td>
</tr>
<tr>
<td>PPV</td>
<td>1299/1310</td>
<td>99 (98–99)</td>
</tr>
</tbody>
</table>

* Using ‘pale or very pale’ as indicator of anaemia.
† Using ‘pale or very pale’ as indicator of severe anaemia.
‡ Using ‘pale or very pale’ as indicator of very severe anaemia.
95% confidence intervals are given in parentheses.

Part 2

The sample was composed of 1529 women (mean age 26 years, range 14–50 years) attending the eight peripheral health units to receive antenatal care. According to the gold standard, the mean Hb was 8.4 g/dl (range 2.9–14.4 g/dl); 1268 women (83%) were classified anaemic and 307 (20%) severely anaemic. Each of the 13 health workers trained on the use of HCS performed on average 109 readings with HCS (range 23–198).

More than 70% of the Hb estimation made by HCS were within 1 g/dl from the gold standard and 89% were within 2 g/dl. Figure 1 presents the differences from the gold standard in gram per decilitre obtained by each reader. The linear regression correlation coefficient of the 13 readers ranged from 0.40 to 0.92 with values <0.50 for three readers, between 0.50 and 0.80 for five readers and >0.80 for five readers.

Figure 2 presents the distribution of sensitivity and specificity obtained by using HCS and CSA methods for anaemia (a) and severe anaemia (b). According to the classification proposed, 46% (six of 13) of the health workers performed excellently and 54% (seven of 13) adequately in the identification of anaemia using HCS, while no one performed excellently and only 12% (one of eight) performed adequately using CSA. In the identification of severe anaemia with HCS, 45% (five of 11) of the health workers performed excellently and 45% (five of 11) adequately, one health worker (9%) obtained poor performance with HCS (in two cases there were no severely anaemic individuals in the sample at the clinic, and therefore it was not possible to evaluate the performance of the workers). No one performed excellently and 50% (four of eight) performed adequately using the CSA. A total of 467 individuals with Hb >7 mg/dl were erroneously identified as severely anaemic by CSA (184 in the case of HCS).

In each health unit, the health workers using HCS did consistently better in identifying anaemic individuals than those who used clinical examination. A better performance in identifying severely anaemic individuals was...
obtained by HCS in 63% (seven of 11) of the cases while a similar performance was reached by CSA in 36% (four of 11). The performance of workers using HCS was never worse than that of those who used CSA. The performances of clinical identification, when only 'very pale' was used as indicator of severe anaemia or very severe anaemia were worse than when 'pale or very pale' was used (data not presented).

Training
For more than 90% of the personnel, a 2-day training session with appropriate material is sufficient to obtain adequate or excellent performances with CSA.

Discussion
Identification of anaemic individuals
The apparently poor specificity of HCS in identifying anaemic individuals was due to individuals with Hb values between 11 and 12 g/dl (86 individuals in part 1 and 90 in part 2), who were classified as 10 g/dl (which is the correct classification according to the HCS instruction). This classification of a group of individuals as anaemic, who were, in reality, slightly over the threshold, lowered the specificity of the HCS. If the group of individuals with Hb between 11 and 12 g/dl is removed from calculations, the specificity reaches 92%. The performance of CSA in identifying anaemic individuals was unsatisfactory in both parts of the study with 1225 anaemics not being identified of a total of 2538.

Identification of severely anaemic individuals
The Haemoglobin Colour Scale has the capacity to correctly identify severely anaemic individuals with an accuracy similar to that observed at higher Hb levels. The sensitivity of HCS for severe anaemia was slightly (but significantly) superior to CSA. In our opinion, clinical evaluation should be used where no alternative methods are available, but because of its low specificity, its use could result in the classification of a large number of normal and anaemic individuals as severely anaemic.

Identification of very severely anaemic individuals
Of the 141 very severely anaemic individuals in the sample, 130 were identified by HCS and 137 by clinical investigation. This similar sensitivity of the two methods, however, should be judged in the light of the PPV. HCS incorrectly classified 133 individuals with Hb ≥ 5 mg/dl as severely anaemic, while CSA incorrectly classified 661 individuals. This difference would have implied different repercussions in terms of number of people to refer. In addition, in the group of individuals that were incorrectly classified as very severely anaemic by HCS, the mean Hb was 6 g/dl; 63% had an Hb > 6 g/dl.
One of the limitations of CSA is that the method works only as a binary indicator (normal and pale). HCS, on the contrary, allows at least five different classifications for each individual. Therefore, even individuals who are not exactly classified by HCS are frequently included in a neighbouring category. We tried to overcome this limitation by introducing the possibility of classifying pallor as ‘very pale’. However, in this study, this does not increase the performance of CSA in the identification of anaemic and very severe anaemic individuals.

As indicated by Paddle (2002), these results confirm that the performance of HCS is not as good as a digital haemoglobinometer. However, in areas where more sophisticated instruments are not available, HCS, at a cost of around 0.03 US$ per test, could enhance the performance of health workers in the diagnosis and treatment of anaemia and severe anaemia and in the identification and referral of very severely anaemic individuals.

Acknowledgements

We thank the Ministry of Health of Zanzibar, United Republic of Tanzania and the personnel of the Pemba Helminth Control Programme for the field-testing. The study was financially supported by WHO/DCT (Devices and Clinical Technology team). Further information on the HCS can be obtained at hbcolourscale@who.int.

References


---

**Authors**

Dr Antonio Montresor (corresponding author) and Dr Lorenzo Savioli, Communicable Diseases Control, Prevention and Eradication, World Health Organization, 1211 Geneva 27, Switzerland. Fax: +41 22 791 4869; E-mail: montresora@who.int, saviolil@who.int

Mahdi Ramsan and Nassor Khalfan, Pemba Public Health Laboratory Ivo de Carneri, Chake Chake, Pemba, Zanzibar. E-mail: fondcar@zanzinet.com

Dr Marco Albonico, Ivo de Carneri Foundation, 20127 Milan, Italy. E-mail: albonico@tin.it

Rebecca J. Stoltzfus, Division of Nutritional Science, Cornell University, Ithaca, NY 14853, USA. E-mail: rjsb2@cornell.edu

James M. Tielsch, Center for Human Nutrition, Department of International Health, Bloomberg School of Public Health, Baltimore, MD 21205, USA. E-mail: jtielsch@jhsp.edu