Diagnostic accuracy comparison between clinical signs and hemoglobin color scale as screening methods in the diagnosis of anemia in children

Comparação da acurácia do diagnóstico dos sinais clínicos e da escala de cores de hemoglobina como métodos de triagem no diagnóstico da anemia em crianças

Abstract

Objectives: to compare the validity and reproducibility of clinical signs with the World Health Organization hemoglobin color scale.

Methods: Two hundred six children in the age range of 6-23 months, at the Instituto Materno Infantil Prof. Fernando Figueira, IMIP, were assessed. Two examiners evaluated the clinical signs and the hemoglobin color scale of each child at the different times. The hemoglobin value was used as a standard for validation.

Results: in more than 90% of cases the agreement between the values of the color scale and the laboratorial hemoglobin was <2 g/dL. Between the clinical signs the highest sensitivity level for diagnosing Hb<11 g/dL was presented by the hemoglobin color scale (75.7%). For moderate/severe anemia Hb<9g/dL the highest sensitivity was shown by combined palmar or conjunctival pallor (74.3%) and by the color scale (52.5%), according to the first and second observer, respectively. The highest specificity level for Hb<11 g/dL was presented by palmar pallor in comparison with the mother’s palm and conjunctival pallor (100%). For Hb<9 g/dL the highest specificity was presented by the hemoglobin color scale (91.9%).

Conclusions: this study suggests that moderate/severe anemia can be diagnosed either by clinical signs or by the color scale, while, in cases of mild anemia, the better diagnosis tool appears to be the color scale.

Key words Anemia, Hemoglobins, Diagnosis, Child

Resumo

Objetivos: comparar a validade e reprodutibilidade dos sinais clínicos (palidez palmar e conjuntival) com a escala de cores da hemoglobina da Organização Mundial de Saúde.

Métodos: pesquisa realizada em 206 crianças de 6-23 meses, atendidas nos ambulatórios de puericultura e pediatria do Instituto Materno Infantil Prof. Fernando Figueira, IMIP. Os sinais clínicos e a escala de cores da hemoglobina foram avaliados por dois observadores em diferentes momentos. A validação foi realizada utilizando-se a hemoglobina como padrão.

Resultados: em mais de 90% dos casos, a concordância entre os valores da escala de cores e da hemoglobina laboratorial encontravam-se em <2 g/dL. Entre os sinais clínicos e a escala de cores, a maior sensibilidade para diagnosticar anemia (Hb<11 g/dL) foi apresentada pela escala de cores da hemoglobina (75,7%) e na anemia moderada/grave (Hb<9 g/dL), pela palidez palmar ou de conjuntiva ocular combinadas (74,3%) e pela escala de cores (52,5%), para o primeiro e segundo examinadores, respectivamente. A maior especificidade foi apresentada na palidez palmar comparada à palma da mãe e palidez de conjuntiva (100%) para Hb<11 g/dL e a escala de cores da hemoglobina (91,9%) para a Hb<9 g/dL.

Conclusões: o presente estudo sugere que o diagnóstico da anemia moderada/grave seja possível através dos sinais clínicos e da escala de cores e, no caso da anemia leve, o melhor instrumento parece ser a escala de cores.

Palavras-chave Anemia, Hemoglobinas, Diagnóstico, Criança
Introduction

Iron deficiency anemia is the most prevailing nutritional deficiency on a global scale, especially in developing countries. The World Health Organization (WHO) estimates that more than two billion people are anemic, which is about one-third of the world population. It is also estimated that, for each anemic person, at least one other suffers from iron deficiency.

The prevalence of anemia among under-four-year old children is as high as 39% in non-industrialized countries, but industrialized countries also present a significant prevalence rate, as high as 20.1%.1

Children, especially young ones, and women of reproductive age, particularly during pregnancy, are more vulnerable to anemia due to their unique physiological characteristics. Typical consequences of iron deficiency anemia include decreased development and motor coordination, decreased language and cognitive capacity development, psychological and behavioral disturbances, and decreased physical activity. Hemoglobin analysis is one simple and most economical laboratory parameter and is thus used quite frequently in population studies.

Another method for diagnosing anemia, studied and used in countries with regions where no laboratories or few available health resources are present, is the physical exam, which aims to identify skin or mucous membrane pallor as clinical signs of anemia. Anatomical segments such as the conjunctiva, palm, nail bed, lips and tongue have been used to identify anemia.

Due to the need for intervention in high infant morbidity and mortality rates, especially in developing countries, the Pan-American Health Organization (PAHO) World Health Organization and the United Nations International Children's Emergency Fund (UNICEF) formulated Guidelines for Integrated Management of Childhood Illness (IMCI), which guide the evaluation and classification of anemia in children based on the clinical signs of light or severe palmar pallor.

Moving beyond clinical signs in diagnosing anemia, Stott and Lewis developed a new and alternative method, namely the WHO hemoglobin color scale. Initially, the technique involved the comparison of a blood drop collected on a piece of special filter paper with a color scale including ten shades of red. Later on, scales were developed with six shades of red and various studies have been carried out to verify their performance.

Since IMCI guidelines are being widely disseminated in Brazil Northeast, this study pursues understanding the new WHO method comparing clinical signs validity and reproducibility (palmar and conjunctival pallor) with the WHO hemoglobin color scale when diagnosing anemia in children between 6 and 23 months of age.

Methods

This study was developed at the Instituto Materno Infantil Prof. Fernando Figueira (IMIP), a public hospital in Recife, in the state of Pernambuco, in the Northeast of Brazil. IMIP has been accredited as a teaching unit by the Ministry of Education and as a National Reference Center by the Ministry of Health. Data were collected at the child and pediatric units.

A sample of 206 children was selected to compare anemia diagnosis through clinical signs such as palmar and conjunctival pallor to the WHO hemoglobin color scale, method during August-October 2001. Sample size was based on the formula for estimating the sensitivity and specificity of diagnostic tests.

N = \frac{Z \times Z (P (1-P))}{(D^2/D)}

D was determined (width of the confidence interval) to be 5%, and, hence, \( z = 1.96 \). To compensate for losses, N (sample size) was increased by 5% to 204 children.

The \( p \) value was chosen based on the highest sensitivity criterion (85.2%) from the study of Montresor et al., for anemia diagnosis, using the WHO hemoglobin color scale.

The study involved those children between 6 and 23 months of age who were accompanied by their biological mother during their medical visits to the pediatric unit. Children suffering from chronic diseases (cardiopathies, rheumatic diseases, nephropathies, gastrointestinal disorders, respiratory diseases, etc.), whose diagnosis had been confirmed at the time of the medical visit, were excluded.

Clinical signs were evaluated by two examiners, one nurse and one pediatrician, previously trained on IMCI guidelines in the clinical evaluation of children’s palmar and conjunctival pallor under artificial light (a fluorescent lamp). Each of the examiners observed the children’s conjunctiva and hand palm at different times, evaluating pallor on the basis of personal judgment. Following, the biological mother’s hand palm was evaluated as a standard to identify palmar pallor.
To evaluate palmar pallor, the examiner opened one of the children's hands, partially extending the fingers, and observed the palm color.

Clinical evaluation of conjunctival pallor was carried out by means of the mucous membrane discoloration concept, that is, the decrease or loss of the reddish pink color. To evaluate conjunctival pallor, the examiner turned the lower eyelid and directly observed the color.

The same examiners also evaluated the WHO hemoglobin color scale. They received two days of training prior to use. Training was performed through blood samples, selected at IMIP laboratory, containing ethylene diamine tetra-acetic acid (EDTA), including a demonstration of the method, after which the examiners compared the hemoglobin lab results with the values obtained on the color scale. Twenty blood samples were used for each examiner. The technique for using this scale determines the use of a fresh blood drop on a special filter paper of the exact size to cover the opening on the scale (diameter of approximately 1 cm). Evaluation of the hemoglobin color scale was carried out 30 seconds following blood collection, based on the separate judgment of each examiner, comparing the blood drop to scale colors. Hemoglobin concentration was determined in accordance with the values on the scale, ranging from 4 to 14 g/dL, including grades of 4, 6, 8, 10, 12 and 14 g/dL.

Laboratory analysis of hemoglobin levels was only performed in the same day after the examiner's evaluation of the clinical signs and of the color scale. The mothers were not informed of the diagnosis.

The project was approved by the Committee for Ethics and Research on Human Beings of the National Ethics Committee. All children who were diagnosed as anemic received oral iron sulphate treatment.

Collected data were processed by the Epi-info, version 6.04 (Centers for Disease Control and Prevention, Atlanta, GA, USA) and Statistica (Stata Soft Inc, Tulsa, OK, USA), computer programs. Anemia diagnosis was based on standards set by the World Health Organization, which accepts hemoglobin concentration values of <11 g/dL\(^1\) as a parameter to confirm the diagnosis of anemia in children between 6 and 59 months of age, and <9 g/dL\(^2\) for moderate/severe anemia.

The validation (sensitivity, specificity, positive predictive value and negative predictive value) was carried out by using hemoglobin value as a reference standard,\(^12\) determined by a Coulter T-890 automatic cell counter, based on vein blood collection.

Bland and Altman\(^14\) plots were used to analyze the agreement between the color scale evaluation and laboratorial hemoglobin levels.

**Results**

The prevalence of anemia in the children of this sample was 92.7%, with a hemoglobin average of 9.6 ± 1 g/dL. Prevalence of moderate/severe anemia was 20.9%.

The two examiners found prevalences of 69.1% and 73.7% by the WHO hemoglobin color scale. The

---

**Figure 1**

Bland-Altman plots for hemoglobin color scale compared with hemoglobin reference.
and 23 months. Gender distribution was 51.2% males and 48.8% females.

Correlation between color scale and laboratorial hemoglobin levels was studied by Bland-Altman plots (Figure 1). For the first examiner, in 67.3% of cases the difference between hemoglobin values of the color scale and the laboratory was within the interval of ± 1 g/dL, and in 92.1% was ± 2 g/dL. For the second examiner, the difference in 66% of cases was within the interval of ± 1 g/dL, and in 92.6% was ± 2 g/dL.

Data analysis revealed greater sensitivity when combining palmar with conjunctival pallor, followed by palmar pallor alone for diagnosing anemia in children's median age was eight months old, with 67.3% of them between 6 and 11 months, 25.6% between 12 and 17 months and 7.1% between 18 and 23 months. Gender distribution was 51.2% males and 48.8% females.

Correlation between color scale and laboratorial hemoglobin levels was studied by Bland-Altman plots (Figure 1). For the first examiner, in 67.3% of cases the difference between hemoglobin values of the color scale and the laboratory was within the interval of ± 1 g/dL, and in 92.1% was ± 2 g/dL. For the second examiner, the difference in 66% of cases was within the interval of ± 1 g/dL, and in 92.6% was ± 2 g/dL.

Data analysis revealed greater sensitivity when combining palmar with conjunctival pallor, followed by palmar pallor alone for diagnosing anemia in

<table>
<thead>
<tr>
<th>Clinical Signs</th>
<th>Sensitivity % (95%CI)</th>
<th>Specificity% (95%CI)</th>
<th>Positive Predictive Value % (95%CI)</th>
<th>Negative Predictive Value % (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examiner I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palmar pallor</td>
<td>40.3 (33.4 - 47.7)</td>
<td>66.7 (38.7 - 87.0)</td>
<td>93.9 (85.7 - 97.7)</td>
<td>8.1 (4.2 - 14.7)</td>
</tr>
<tr>
<td>Palmar pallor in comparison with the mother's palm</td>
<td>31.4 (25.0 - 38.6)</td>
<td>86.7 (58.4 - 97.7)</td>
<td>96.8 (87.8 - 99.4)</td>
<td>9.0 (5.1 - 15.2)</td>
</tr>
<tr>
<td>Conjunctival pallor</td>
<td>20.6 (15.1 - 27.3)</td>
<td>92.3 (62.1 - 99.6)</td>
<td>97.4 (84.6 - 99.9)</td>
<td>7.7 (4.2 - 13.4)</td>
</tr>
<tr>
<td>Palmar or conjunctival pallor</td>
<td>49.4 (42.0 - 57.0)</td>
<td>76.9 (46.0 - 93.8)</td>
<td>96.7 (90.1 - 99.2)</td>
<td>9.9 (5.1 - 17.9)</td>
</tr>
<tr>
<td>Colour scale</td>
<td>72.2 (65.1 - 78.4)</td>
<td>73.3 (44.8 - 91.1)</td>
<td>97.1 (92.3 - 99.1)</td>
<td>17.5 (9.4 - 29.5)</td>
</tr>
<tr>
<td>Examiner II</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palmar pallor</td>
<td>28.3 (22.0 - 35.4)</td>
<td>93.3 (66.0 - 99.7)</td>
<td>98.1 (88.6 - 99.9)</td>
<td>9.6 (5.5 - 15.9)</td>
</tr>
<tr>
<td>Palmar pallor in comparison with the mother's palm</td>
<td>15.3 (10.6 - 21.5)</td>
<td>100.0 (74.7 - 100.0)</td>
<td>100.0 (85.0 - 100.0)</td>
<td>8.8 (5.2 - 14.4)</td>
</tr>
<tr>
<td>Conjunctival pallor</td>
<td>24.8 (18.5 - 32.5)</td>
<td>100.0 (71.7 - 100.0)</td>
<td>100.0 (88.8 - 100.0)</td>
<td>9.9 (5.6 - 16.7)</td>
</tr>
<tr>
<td>Palmar or conjunctival pallor</td>
<td>33.1 (25.9 - 41.1)</td>
<td>92.3 (62.1 - 99.6)</td>
<td>98.1 (88.6 - 99.9)</td>
<td>10.3 (5.6 - 17.6)</td>
</tr>
<tr>
<td>Colour scale</td>
<td>75.7 (68.5 - 81.8)</td>
<td>53.3 (27.4 - 77.7)</td>
<td>94.9 (89.4 - 97.8)</td>
<td>16.0 (7.6 - 29.7)</td>
</tr>
</tbody>
</table>
Comparison between clinical signs and hemoglobin scale methods

Specificity displayed higher rates for conjunctival pallor and for palmar pallor compared to the mother's hand palm. NPV were high, exceeding 80% for all signs and for the WHO hemoglobin color scale, while PPV were higher for the hemoglobin color scale.

Discussion

The high anemia prevalence (92.7%) found in our study in the age range between 6 and 23 months is consistent with WHO's report that in the first five years of life, iron deficiency is highest in the second children. Conjunctival pallor obtained the highest specificity (Table 1).

When evaluating the capacity of the WHO hemoglobin color scale to identify anemic and non-anemic individuals, a sensitivity above 72% and specificity rates above 53% was determined.

Table 1 displays positive predictive values (PPV) above 90.0% for all signs and for the hemoglobin color scale. The highest negative predictive value (NPV) went to the hemoglobin color scale.

The combination of palmar or conjunctival pallor and isolated palmar pallor (Table 2) reached the highest sensitivity coefficients for diagnosing moderate/severe anemia.

Table 2

Sensibility, specificity, positive and negative predictive values of clinical signs and the World Health Organization color scale to diagnose anemia moderate/severe in children, using hemoglobin as the gold standard, according to examiner I and II.

<table>
<thead>
<tr>
<th>Clinical Signs</th>
<th>Sensitivity % (95%CI)</th>
<th>Specificity% (95%CI)</th>
<th>Positive Predictive Value % (95%CI)</th>
<th>Negative Predictive Value % (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examiner I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palmar pallor</td>
<td>60.5 (44.4 - 74.6)</td>
<td>65.6 (57.8 - 72.8)</td>
<td>31.7 (22.1 - 43.0)</td>
<td>86.3 (78.7 - 91.6)</td>
</tr>
<tr>
<td>Palmar pallor in comparison</td>
<td>48.8 (33.6 – 64.3)</td>
<td>74.8 (67.3 – 81.2)</td>
<td>33.9 (22.6 – 47.1)</td>
<td>84.7 (77.6 – 90.0)</td>
</tr>
<tr>
<td>with the mother's palm</td>
<td>34.3 (19.7 – 52.3)</td>
<td>83.5 (76.6 – 88.8)</td>
<td>31.6 (18.0 – 48.8)</td>
<td>85.2 (78.4 – 90.2)</td>
</tr>
<tr>
<td>Conjunctival pallor</td>
<td>74.3 (56.4 – 86.9)</td>
<td>58.2 (50.1 – 65.9)</td>
<td>28.3 (19.6 – 38.8)</td>
<td>91.1 (83.3 – 95.6)</td>
</tr>
<tr>
<td>Palmar or conjunctival pallor</td>
<td>45.2 (30.2 – 61.2)</td>
<td>91.9 (86.2 – 95.4)</td>
<td>59.4 (40.8 – 75.8)</td>
<td>86.5 (80.2 – 91.1)</td>
</tr>
<tr>
<td>Colour scale</td>
<td>41.9 (27.4 – 57.8)</td>
<td>77.6 (70.1 – 83.7)</td>
<td>34.0 (21.9 – 48.4)</td>
<td>82.9 (75.6 – 88.4)</td>
</tr>
<tr>
<td>Palmar pallor in comparison</td>
<td>21.4 (10.8 – 37.2)</td>
<td>87.8 (81.4 – 92.3)</td>
<td>32.1 (16.6 – 52.4)</td>
<td>80.6 (73.7 – 86.1)</td>
</tr>
<tr>
<td>with the mother's palm</td>
<td>30.3 (16.2 – 48.9)</td>
<td>78.8 (70.9 – 85.2)</td>
<td>25.6 (13.6 – 42.4)</td>
<td>82.4 (74.6 – 88.3)</td>
</tr>
<tr>
<td>Palmar or conjunctival pallor</td>
<td>42.4 (26.0 – 60.6)</td>
<td>71.5 (63.1 – 78.7)</td>
<td>26.4 (15.7 – 40.6)</td>
<td>83.8 (75.5 – 89.7)</td>
</tr>
<tr>
<td>Colour scale</td>
<td>52.5 (36.3 – 68.2)</td>
<td>89.2 (82.8 – 93.5)</td>
<td>56.8 (39.6 – 72.5)</td>
<td>87.4 (80.8 – 92.1)</td>
</tr>
</tbody>
</table>

year. This is due to both a low dietary iron consumption and rapid growth in the first year. It's possible that in Brazilian Northeast, local dietary practices associated to poor socio-economic conditions contribute to this scenario.

In validation studies, disease prevalence may modify the instrument performance indicators, showing high PPV for high prevalence or high NPV for low prevalence. Therefore, the test ability to confirm anaemia (PPV >90% for all signals and for the hemoglobin colour scale) for self-diagnosed subjects, as well as the test ability to identify high NPV (>80% for all signals and for the hemoglobin colour scale), should be regarded as strictly valid only for the population under study in this work.

The use of the WHO hemoglobin color scale has been studied with lab blood samples and has displayed a high correlation level with the concentration of blood hemoglobin in adults when using the linear correlation method. Montresor et al., have shown a 0.80 correlation rate between the color scale and the hemoglobin values measured by the HemoCue in pre-school children.

In this study, the use of the Bland-Altman plot to analyze the correlation between two methods of clinical measuring, emphasized that over 90% of the agreements between hemoglobin values of the color scale and laboratory values were comprised in the interval of ± 2 g/dL, which was accepted considering the scale's color intervals.

The WHO hemoglobin color scale, with sensitivity above 70%, revealed to be a better method than clinical signs to diagnose anemia in children. Using an instrument with pre-established color standards reduces error subjectivity caused by individual interpretation, like in the case of cutaneous-mucous pallor. However, specificity was lower than for most of the clinical signs. The clinical signs displayed a higher amount of false negative results as compared to the hemoglobin color scale. Therefore, it is difficult to identify skin and mucous pallor in cases of mild anemia, increasing the risk of not diagnosing the children who are actually ill, meaning that many children diagnosed as non-anemic were actually anemic.

The results described above corroborate the studies by Ingram and Ingram and Lewis and Montresor et al., in which sensitivity for diagnosing mild anemia was higher for the color scale than for clinical signs.

The scale also revealed to be a better instrument for diagnosing moderate/severe anemia, according to the second observation of this study, which also corroborates the studies by Ingram and Lewis and Montresor et al., when considering cut-off points of 8-9 g/dL and <7 g/dL, respectively.

However, for most of the clinical signs, except for conjunctival pallor, the first examiner displayed a greater capacity to identify true positive results in cases of moderate/severe anemia, increasing the test sensitivity. In spite of using different cut-off points for severe anaemia, Ingram and Lewis and Montresor et al., showed that there are no errors in diagnosing severe anemia by means of clinical signs, considering cut-off points <6 g/dL and <5 g/dL, respectively. However, higher hemoglobin values are hard to detect and interpret through clinical signs.

For anemia diagnoses, the ideal situation would be to use indices resulting in a higher incidence of false positive than false negative results, since the treatment of non or mildly anemic children who are diagnosed as moderately or severely anemic would not be a problem for most of them. However, an increase in false negative results would imply that anemic children were not included in anemia combat programs, which would be very harmful in places with high anemia prevalence rates.

Various authors have also reported on the performance of the scale, verifying sensitivities above 80%, .

One of the advantages in using the hemoglobin color scale is that the examiners can be trained in an easy, fast and economically efficient way, with a minimal need of material, to distinguish between the different color variations of the scale, without observational bias. However, in this training, it is important to follow the instructions on how to hold the instrument and use the correct thickness of the blood drop, adequate environmental luminosity and the established reading time.

Other advantages of the color scale are its low cost in comparison with other methods to determine hemoglobin concentration, and durability, since instrument colors remain's stable during various months, even when in constant use.

It's important to note that in laboratory diagnoses, hemoglobin reference values are expressed in 0.1 g/dL, while the scale is read at 2 g/dL, or 1 g/dL intervals when the examiner considers the value may lie between two shades. According to the guidelines described in the instructions for color scale usage, in case of doubt, the examiner should consider the lowest value, thus increasing the number of false positive results, which would be of interest to communities with high anemia prevalence. Even when guidelines are followed, confidence intervals specificity for both examiners in this study revealed a wide variation suggesting that color scale use holds
Comparison between clinical signs and hemoglobin scale methods

The use of the WHO hemoglobin color scale for anemia diagnosis in the routine practice of health professionals, not as a substitute for the laboratory test, but as a useful method for early diagnosis of a possible decrease in hemoglobin levels, with the goal of fast intervention and, therefore, prevention of a possible problem.

Acknowledgements

The authors thank the mothers and children who participated of the research and IMIP for the support in the field work. This study was funded by the Pan-American Health Organization and Brazil’s Ministry of Health.

References

3. INACG Symposium; 1999 March 12; Durban, South Africa. Washington (DC); 2000.