Haemoglobin colour scale for anaemia diagnosis where there is no laboratory: a systematic review

Julia Critchley* and Imelda Bates

Accepted 31 August 2005

Background
Anaemia is a major public health problem, in poor countries most of the cases are diagnosed clinically. This is inaccurate and the haemoglobin colour scale (HCS) has been developed as an inexpensive, simple alternative for assessing anaemia. Laboratory and community studies have assessed its diagnostic accuracy, but controversy over its validity and usefulness remains. We carried out a systematic review to identify and summarize studies, explain heterogeneity, and make recommendations for future research.

Methods
We searched electronic databases (MEDLINE, EMBASE, CINAHL, and Science Citation Index), checked documents and references, and contacted experts. We included all the studies comparing HCS diagnostic accuracy with a reference standard. Both reviewers independently screened titles and abstracts, assessed studies for inclusion, appraised quality, and extracted data.

Results
We included 14 studies, mostly from sub-Saharan Africa. Studies had heterogeneous populations, health care settings, anaemia prevalence, and findings. HCS sensitivity for detecting anaemia was high in most of the studies (75–97%); specificity was generally lower (41–98%). Sensitivity and specificity were higher for laboratory-based studies compared with more pragmatic ‘real-life’ studies, and the ‘study setting’ appeared to explain some of the heterogeneity. Five studies compared the HCS with clinical diagnosis; sensitivity was higher for the HCS in four studies, but specificity was often higher with clinical diagnosis. A few studies evaluated the HCS in situations where there was no laboratory.

Conclusions
The HCS may improve anaemia diagnosis where there is no laboratory, but there is a need for policy-relevant diagnostic research which is pragmatic, implementation-focused and assesses clinical outcomes. This requires a different approach and research skill-mix from efficacy studies.

Keywords
Anaemia, haemoglobin, haemoglobinometry, developing countries, diagnosis, health plan implementation

Introduction
Anaemia is the world’s second leading cause of disability1 and one of the most serious global public health problems. Anaemia affects over half of the pre-school children and pregnant women in developing countries.2 In poorer malaria endemic countries it is one of the commonest preventable causes of death in children <5 years of age and in pregnant women.3 The HIV pandemic is worsening the severity and increasing the prevalence of anaemia.4 Even mild and moderate anaemia has serious clinical consequences. In pregnancy it is associated with increases in stillbirths and low birth weight babies, and in children it causes stunting and impaired intellectual development.5

Diagnosis of anaemia is problematic in resource-poor settings. Clinical diagnosis is the most common method of detecting anaemia but this is inaccurate. For example, it is only 65% sensitive for detecting haemoglobin levels of <7 g/dl in pregnant women in Kenya6 and is even less useful for detecting non-severe anaemia. Most of the existing methods for measuring haemoglobin are expensive, complex, or impractical for use.

Liverpool School of Tropical Medicine, Liverpool, UK
* Corresponding author. International Health Research Group, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK. E-mail: juliac@liverpool.ac.uk
in primary care in poor countries. The ‘gold standard’ haemoglobin-cyanide method uses inexpensive but toxic consumables, depends on accurate sample dilutions, and requires regular supervision and careful quality control processes. Other methods, such as the HemoCue® system, are simple to operate but use costly disposable cuvettes making them too expensive for widespread use. A cheap reliable means of assessing haemoglobin levels is therefore required for use in places where there is no laboratory.

Development and use of the haemoglobin colour scale

The haemoglobin colour scale (HCS), a simple, rapid and cheap method for estimating haemoglobin concentration with a finger prick blood sample, has been developed for use in resource-poor settings where there is no laboratory. The method relies on comparing the colour of a drop of blood absorbed onto a filter paper with standard colours on a laminated card, varying from pink to dark red. These colours correspond to haemoglobin levels of 4, 6, 8, 10, 12, and 14 g/dl. Intermediate shades can be identified, allowing haemoglobin levels to be judged to 1 g/dl. This is the only method currently available that is affordable for widespread primary care use by the most impoverished countries.

A diagnostic test for assessing anaemia in communities and primary care facilities in resource-poor settings should be inexpensive, rapid, simple to perform on a finger prick sample, and reasonably accurate. Because the HCS is divided into increments it is not surprising that studies have shown it to be less accurate than the laboratory methods of measuring haemoglobin [such as the hemoglobin-cyanide (HCN) and HemoCue® methods] which measure haemoglobin levels to the nearest 0.1 g/dl. Interpretation of HCS results is highly variable depending on the context. Opinions vary from the recognition that the HCS is accurate enough to be useful for anaemia diagnosis in resource-poor settings to suggestions that its poor accuracy ‘renders its use questionable’. These opinions partly reflect the wide range of results from studies of the accuracy of the tests in different contexts, and a lack of clarity in defining whether these studies were designed to assess accuracy or usefulness of the HCS in practice. To provide an overview of the evidence available to support widespread introduction of the HCS we have performed the first systematic review to assess the objectives, quality, and results of all published studies evaluating the performance of the HCS compared with reference methods and clinical diagnosis. We have examined possible explanations for heterogeneity between studies, and developed recommendations for research strategies assessing the practical usefulness of such diagnostic tests in resource-poor settings.

Methods

Types of studies included

We included any study comparing the HCS with a ‘gold standard’ method of measuring anaemia, such as haemoglobin-cyanide (cyanmethaemoglobin method), in a population-based or hospital-based sample. For pragmatic reasons, we also included HemoCue® as a ‘gold standard’ method of measuring anaemia in poorer tropical countries.

Identification and selection of studies

Electronic databases, MEDLINE, EMBASE, CINAHL, and Science Citation Index (up to May 2005), were searched using keywords ‘Haemoglobin Colour Scale’ and alternative spellings (hemoglobin or color) and the full text of any articles that appeared relevant was retrieved. Both authors assessed studies for inclusion, using a pre-designed, piloted inclusion form. We also searched for relevant studies by screening titles and abstracts listed in the WHO Haemoglobin Colour Scale bibliography, by checking lists of references in relevant studies (all included studies and reviews), and by personal communication.

Data extraction and assessment of methodological quality

Both reviewers’ extracted data using a pre-designed form. This included information on whether the studies were designed to assess accuracy or usefulness in practice; study characteristics, such as country and location; definitions of anaemia and anaemia prevalence; reference (‘gold standard’) method of measuring anaemia; health care setting and workers; any training received, including the length and type of training; sample size; statistical methods; and results.

Studies were categorized by both reviewers as essentially ‘laboratory based’ (efficacy studies), ‘real-life’ (effectiveness studies), or ‘unclear’ on the basis of two criteria: (i) whether the HCS was tested using samples from patients in clinic settings (real-life), or samples of known haemoglobin from laboratories (laboratory based) and (ii) level of training and supervision received by those using the HCS (in ‘real-life’ studies users received not more than 1 h of training, more extensive training and constant supervision were classified as laboratory based). Studies classified as ‘unclear’ contained elements of both the study types (such as real-life but with extensive training), or did not report sufficient details in order to make a judgement.

The data extraction form was designed by consulting accepted guidelines and empirical evidence for the importance of various aspects of the design of diagnostic studies, and was piloted on two studies before use. We contacted authors for more details where specific items [such as 95% confidence intervals (95% CIs)] were not reported. Methodological quality of studies was assessed by both reviewers, adapting validated guidelines specifically developed for assessing the quality of diagnostic studies. The included items are as follows:

- Clear descriptions of selection criteria, methods of testing, and the reference standard (‘gold standard’) used.
- Whether blood for the reference standard and HCS were collected and analysed at the same time.
- Whether the HCS test results were interpreted without knowledge of the results of the reference standard, and the results of clinical examination if applicable.
- Whether the same clinical data were available when HCS results were interpreted as would be available when the test is used in practice.
- Whether ‘intermediate’ HCS results were reported.
- Explanation of withdrawals from the study population.
- Whether results from all the health workers were included in the final analysis.
heterogeneity between studies we did not combine them.

Data analysis

We planned to pool estimates of sensitivity and specificity of the HCS across studies wherever appropriate, stratified by ‘real-life’ or ‘laboratory’ based. However, statistical methods for combining results of diagnostic tests are at an early stage of development. In particular, it would be inappropriate to combine estimates of sensitivity and specificity if there is a marked heterogeneity in these estimates, if the disease spectrum is different, or if the diagnostic cut-offs for anaemia are varying between studies. Due to the clinical and statistical heterogeneity between studies we did not combine them statistically, except for the laboratory-based studies with similar findings.

Sensitivity and specificity were stratified by the study type. Diagnostic odds ratios (DORs—ratios of odds of a ‘positive’ HCS result in patients with anaemia compared with a patient without anaemia, the larger the value the stronger the diagnostic evidence) were calculated as an overall measure of test accuracy (see the web version of Table 1, available at IJE Online).

Many studies did not provide any estimate of precision for these statistics (such as a 95% CI) so wherever possible we estimated these from the raw data using StatsDirect, or obtained them from the authors.

Results

Description of studies

Fourteen studies were included (Figure 1 and Table 1). Two studies from the WHO’s HCS bibliography were assessed but excluded because they were pilot studies, primarily concerned with training. Development of the HCS has been an iterative process, and several of the included studies tested prototype versions of the HCS. For example, in one study users found it difficult to differentiate between values of 10 and 12, and the shades were later adjusted. Another study did not allow blood to soak into the paper for 30 s, and Tatsumi29 experimented with altering the amount of blood, suggesting that the method was not optimized.

The objectives, type of patients, and the health care setting in which the HCS was used varied between studies making it difficult to generalize the results. Some studies were carried out under ‘ideal’ settings or in patient populations that are not representative of those presenting to primary care health facilities in developing countries. Six studies were ‘laboratory based’. These assessed the ability of health care workers to obtain accurate results under supervision using the HCS, compared with a reference method. Four studies were ‘real-life’, and were carried out without supervision in antenatal clinics, schools, or rural research clinics after a brief training. The remaining four ‘real-life’ or ‘laboratory based’ contained elements of both, e.g. laboratory testing but with brief training. Use of HCS in a field setting but as part of a randomized controlled trial, and with extensive training.

Most of the studies were conducted in sub-Saharan Africa (Ethiopia, Tanzania, South Africa, Zimbabwe, Malawi, Kenya). Other studies include populations from the UK, Indonesia, India, Honduras, Thailand, and Switzerland. Definitions of anaemia and anaemia prevalence varied between studies. Most of these defined a cut-off for anaemia as a haemoglobin level either <11 g/dl or <12 g/dl. Some studies also assessed the diagnostic accuracy of the HCS at other cut-offs but the definitions of these varied. ‘Severe’ anaemia was generally defined as Hb <7 g/dl or <6 g/dl, but one study also defined a cut-off for ‘very severe’ anaemia as <5 g/dl.

A few studies used different definitions for men and women. Anaemia prevalence varied dramatically, from 5 to 85%. The reference methods used to diagnose anaemia included HemoCue® (7 studies), automated analyser (7), photometer (3) and the copper sulphate method (2). Most of the studies provided training for health care workers using the HCS, though the methodology differed between studies and was usually not clearly described. Several studies also assessed the diagnostic accuracy of other methods, such as clinical examination for anaemia (5 studies), the copper sulphate method (3 studies), the Sahli method (1 study), and the Tallquist paper scale (1 study), against the reference method.

Methodological quality of studies (web Table 2)

In all the studies, health workers using the HCS appear to have been ‘blind’ to results of other methods of measuring haemoglobin. Many studies did not clearly describe their criteria for selecting health workers or patients, or the characteristics or completeness of their sample. This severely hinders interpretation of the results as the sensitivity and specificity of any test may vary in different populations (e.g. by sex, age or disease severity). Several studies used venous blood samples, whereas finger prick samples would be the most likely source of blood if the HCS was used at a community or a primary care level. Two studies used a venous blood sample for the reference standard (automated haematology analyser) but used a finger prick blood sample for the HCS measurement, thereby introducing the possibility of bias.

Some studies excluded results from health workers who were retrospectively shown to have performed poorly on the HCS or allowed health care workers to estimate HCS more than once, without explaining how any discrepancies were resolved. The HCS measures haemoglobin in 2 g/dl increments. Some studies allowed intermediate values (to the nearest 1 g/dl) while others did not, with implications for calculations of diagnostic accuracy. In some studies inappropriate statistics were used to assess the agreement between the HCS and the reference standard (such as correlation coefficients).32

Sensitivity and specificity of HCS for the detection of anaemia

Thirteen studies reported estimates of sensitivity or specificity for anaemia using the HCS and were categorized as ‘real-life’, ‘laboratory based’, or unclear (Figures 2a and b). Estimates of sensitivity and specificity from the five laboratory-based studies were all high, varying from 0.85 to 0.99 for sensitivity and from 0.91 to 1.0 for specificity. All used the same cut-off for diagnosing anaemia (Hb <12 g/dl), so a pooled estimate was calculated from the four studies for which the
<table>
<thead>
<tr>
<th>Study reference (location)</th>
<th>Anaemia prevalence</th>
<th>Patient population</th>
<th>Main results</th>
</tr>
</thead>
</table>
| 23 (Kenya)                | 60%                | 178 children (2–36 months) | Sens: 0.76% (CI 0.68–0.84)   
Spec: 0.44% (CI 0.33–0.56) |
| 20 (Indonesia)            | 38%                | 240 blood donors, 150 samples from the haematology unit | Anaemia: 0.99 (CI 0.96–1.0)   
Severe anaemia: 1.0 (CI 0.40–1.0)   
Spec: Anaemia: 1.0 (CI 0.98–1.0)   
Severe anaemia: 0.99 (CI 0.95–0.99) |
| 24 (Southern Ethiopia)   | 15.1%              | 403 pregnant women | Sens: 0.33–0.44 (CI 0.12–0.78)  
Spec: 0.87–1.0 (CI 0.83–1.0) |
| 10 (Tanzania)             | 85% anaemic, (29% severe, 9% very severe) | 990 children and 643 gynaec/obstetric patients | Sens: Anaemia: 0.97 (CI 0.96–0.97)  
Severe anaemia: 0.92 (CI 0.89–0.94)  
Very severe anaemia: 0.92 (CI 0.84–0.94)  
Spec: Anaemia: 0.41 (CI 0.34–0.47)  
Severe anaemia: 0.86 (CI 0.84–0.88)  
Very severe anaemia: 0.90 (CI 0.88–0.91) |
| 25 (Upper Egypt)          | 17.5%              | 150 children (aged 6–11) | Sens: 0.88 (CI 0.70–0.97)  
Spec: 0.49 (CI 0.40–0.60) |
| 11 (UK)                   | Not relevant       | 408 adult and child inpatients | Not possible to calculate sensitivity/specificity |
| 21 (Multi-centre)         | ~5%                | 2801 blood donors | Sens: 0.85 (CI 0.77–0.91)  
Spec: 0.98 (CI 0.97–0.98) |
| 26 (Tanzania)             | 78.5% anaemia, 3.5% severe anaemia | 614 children (16–83 months) | Sens: Anaemia: 0.85 (CI 0.81–0.88)  
Severe anaemia: 0.74 (CI 0.49–0.90)  
Spec: Anaemia: 0.77 (CI 0.68–0.84)  
Severe Anaemia: 0.99 (CI 0.99–1.00) |
| 17 (South Africa)         | Not reported       | 548 hospital outpatients | Sens: Anaemia: 0.85 (CI 0.81–0.99)  
Severe anaemia: 0.75 (CI 0.81–0.99)  
Spec: Anaemia: 0.77 (CI 0.95–0.98)  
Severe anaemia: 0.99 (CI 0.98–0.99) |
| 27 (South Africa)         | 36%                | 104 adult outpatients | Sens: 0.7  
Spec: 0.85 |
| 28 (Malawi)               | 58.1% (mild) 32% (moderate) 4% (severe) 0.4% (very severe) | 729 pregnant women | Sens: Anaemia: 0.78 (CI 0.74–0.8)  
Moderate: 0.82 (CI 0.77–0.86)  
Severe anaemia: 0.81 (CI 0.65–0.92)  
Very severe: 0.5 (CI 0.12–0.58)  
Spec: Anaemia: 0.50 (CI 0.46–0.55)  
Moderate 0.45 (CI 0.42–0.49)  
Severe anaemia: 0.74 (CI 0.74–0.79)  
Very severe: 0.99% (CI 0.98–0.99) |
true positive, true negative, false negative, and false positive rates could be obtained. This yielded an overall sensitivity of 0.87 (95% CI 0.86–0.89) and specificity of 0.97 (95% CI 0.97–0.98). There was no statistically significant heterogeneity between estimates of sensitivity and specificity (χ² test, P < 0.25 for both).

Estimates of sensitivity and specificity from the ‘real-life’ or unclear studies were lower and more heterogeneous than for the laboratory-based studies, so a pooled estimate was not considered appropriate. Sensitivity for the ‘real-life’ studies varied between 0.76 and 0.88, apart from one outlier (Figure 2a). Sensitivities lower than this may be clinically unacceptable; for example, a sensitivity of 0.75 implies that 25% of the patients with anaemia will not be diagnosed using the HCS. Specificity was lower for the ‘real-life’ studies (Figure 2b), implying that the HCS may falsely diagnose anaemia in many individuals. Although anaemia treatment is cheap and relatively safe, the potential costs of this over-diagnosis, particularly for patients with very limited resources, must be taken into account in assessing the usefulness of the HCS. DORs for anaemia diagnosis were significantly lower (between 2.5 and 7.5) for ‘real-life’ studies, compared with laboratory-based studies (between 58 and 48 000), P = 0.0032 (web Table 1, available at IJE Online; Figure 3).

The haemoglobin cut-off for diagnosing anaemia varied between studies (11–12 g/dl). As the HCS was read only to the nearest 1 or 2 g/dl in these studies we would not necessarily expect such variations to cause heterogeneity between studies, and there was almost no correlation between study-level estimates of sensitivity and specificity (Spearman’s rho = 0.027, P = 0.29), suggesting that factors other than a threshold effect account for most of the heterogeneity. These factors may include characteristics of the population (such as anaemia prevalence and patient group), test operator (level of training and experience), or the test itself (optimal version). For example, one of the ‘real-life’ studies and one of the unclear studies estimated particularly low sensitivity [0.44 (95% CI 0.32–0.56) and 0.23 (95% CI 0.13–0.36), respectively]. However, the former study was carried out in an area of very low anaemia prevalence, and in the latter a prototype version of the HCS scale was

### Table 1 continued

<table>
<thead>
<tr>
<th>Study reference (location)</th>
<th>Anaemia prevalence</th>
<th>Patient population</th>
<th>Main results</th>
</tr>
</thead>
<tbody>
<tr>
<td>29 (Indonesia)</td>
<td>16%</td>
<td>120 Children</td>
<td>Sens: 0.23 (CI 0.13–0.36) Spec: 0.97 (CI 0.89–1.0)</td>
</tr>
<tr>
<td>28 (Multi-site reference centres)</td>
<td>67% (severe 14%)</td>
<td>1213 random blood samples</td>
<td>Sens: Anaemia: 0.86 (CI 0.83–0.89) Severe: 0.94 (CI 0.89–0.97) Spec: Anaemia: 0.91 (CI 0.88–0.93) Severe: 0.98 (CI 0.97–0.99)</td>
</tr>
<tr>
<td>28 (South Africa)</td>
<td>Not relevant</td>
<td>20 blood samples</td>
<td>Inappropriate statistical methods</td>
</tr>
<tr>
<td>28 (South Africa and Thailand)</td>
<td>Not given</td>
<td>113 blood donors</td>
<td>No discrepancies between methods</td>
</tr>
<tr>
<td>29 (South Africa)</td>
<td>Not given</td>
<td>20 blood samples</td>
<td>Sens: Anaemia: 0.89 (CI 0.86–0.93) Severe anaemia: 0.84 (CI 0.77–0.91) Spec: Anaemia: 0.95 (CI 0.91–0.97) Severe anaemia: 0.92 (CI 0.90–0.94)</td>
</tr>
</tbody>
</table>

CI = 95% confidence interval; Sens = sensitivity; Spec = specificity.
used. For both these studies specificity was high (Figure 2b), suggesting that threshold effects may account for some of the heterogeneity, despite the lack of a clear correlation. Similarly, one of the ‘unclear’ studies appeared to have lower specificity, but sensitivity was high\textsuperscript{10} (Figures 2a and b). DORs (Figure 3) appear reasonably consistent within the three sub-groups, suggesting that the study settings account for some of the heterogeneity, except for one study that involved extensive training, where the HCS was more accurate than in any other study.\textsuperscript{20}

Sensitivity and specificity of HCS for the detection of severe anaemia

Six studies detected and reported HCS results for severe anaemia (defined as Hb < 6 or 7).\textsuperscript{10,17,18,20,26,28} Sensitivity was high (>0.84) in most of the studies but in the only ‘real-life’ based study it was just 50% (95% CI 12–58%)\textsuperscript{28} (Figure 2c). If confirmed, sensitivity as low as 50% would be a cause for concern. However, clinical diagnosis is more accurate for severe anaemia compared with mild anaemia, and the HCS is therefore likely to be most useful for detecting mild to moderate cases.

Specificity was very high in all the studies, varying between 86% (95% CI 84–88%) and 99% (95% CI 99–100%) (Figure 2d). Most of the studies only had a small number of patients who were severely anaemic, so larger sample sizes and further studies are required to assess the diagnostic accuracy of HCS for severe anaemia (Figure 4).

Comparison of alternative diagnostic tests for anaemia with HCS

Clinical examination

Five studies reported diagnostic accuracy for anaemia using both HCS and clinical examination\textsuperscript{10,17,24,26,28} (usually by assessing
conjunctival pallor), compared with the same reference standard. In two of the studies the clinical examination was far more intense than would be the case in practice. It involved assessing pallor of nail beds, conjunctiva and palms, and grading the appearance of each site as normal, pale, or very pale.\textsuperscript{10,26} Despite this, in most studies clinical examination performed poorly, especially for detecting mild and moderate levels of anaemia.

Sensitivity varied from 33\% (95\% CI 29–38\%) to 57\% (95\% CI 34–79\%) and specificity from 79\% (95\% CI 73–79\%) to 84\% (95\% CI 79–88\%). The HCS generally had better sensitivity, but not necessarily better specificity, than clinical examination. Only one study reported that clinical examination performed better than the HCS.\textsuperscript{24}

Other tests
One study in Egypt compared the diagnostic accuracy of the ‘Sahli’ method (a subjective method based on visual comparison) and HCS with the HemoCue reference standard.\textsuperscript{25} Sensitivity was similar for both the methods (92.3\%, 95\% CI 75–99\% for Sahli compared with 88.5\%, 95\% CI 70–97\% for HCS), but in both the tests specificity was very low (48.8\%, 95\% CI 40–60\% and 39.0\%, 95\% CI 30.5–48\% for HCS and Sahli, respectively). The authors preferred the HCS as it is quicker to carry out than the Sahli method. Three studies compared HCS with the copper sulphate method for screening blood donors and found that both methods had similar sensitivity and specificity when using HemoCue or haemoglobin cyanide as the reference methods\textsuperscript{20,21,29} (Table 1). One study used results of the HCS, in combination with anaemia symptoms and clinical examination, as part of a scoring system for anaemia, and showed higher sensitivity and specificity than the HCS alone.\textsuperscript{24}

Discussion
We have identified and extracted data from 14 studies to provide an overview of the evidence available about the accuracy and usefulness of the HCS in anaemia diagnosis in laboratory-based and ‘real-life’ settings. The HCS has only been available since 2001, so there are relatively few publications at present (Figure 1). Assessing the methodological quality of these studies
and interpreting their results was difficult due to poor reporting of the study design and methods; different diagnostic criteria for anaemia; and frequently small sample sizes, particularly for severe anaemia. We have identified gaps in the evidence and clear evidence of heterogeneity between study settings. As with any systematic review there is a possibility of bias due to the methods of selecting the publications. However, the search was comprehensive and supplemented by contacting key workers in the field and WHO, and data were extracted independently by two researchers.

Despite these limitations, our review suggests that the HCS may be a useful tool for anaemia diagnosis. Sensitivity and specificity were very high in laboratory-based studies, but deteriorated substantially as the studies moved from accuracy assessments in ‘ideal’ populations in the laboratory to more ‘real-life’, pragmatic evaluations. The most likely explanation for this is the less intensive training and supervision available in the field. Estimates of sensitivity and specificity appeared heterogeneous between studies, but estimates of DORs were reasonably consistent within our laboratory-based or ‘real-life’ sub-groups, suggesting that the study ‘setting’ can account for some of the heterogeneity between included studies. The reduced HCS accuracy in field studies implies that further research is required to define and evaluate its role in primary health care.

At present the most widely used method for detecting anaemia in settings where there is no laboratory is clinical diagnosis, and this is well recognized (confirmed by this review) to be useful only for very marked anaemia. The HCS performed better than clinical diagnosis alone for detecting mild and moderate anaemia, but this advantage is reduced as anaemia becomes more severe and hence more clinically obvious. Even for severe anaemia, the HCS may have the advantage of being semi-quantitative, whilst clinical diagnosis is always qualitative. The HCS is likely to be most useful in areas with moderate to high anaemia prevalence, though there may be a trade-off with lower specificity—in which case the costs and implications of a significant number of false positive results for patients and the health services must be considered. Further studies are therefore needed to assess the benefits and disadvantages of using the HCS in specific patient populations (e.g. ante-natal women and young children) in high prevalence rural areas. The aim of a device such as the HCS should be to improve identification of anaemic individuals, and not to replace clinical examination, so it is important to investigate the most effective way of combining these approaches. One study used a scoring system based on a combination of HCS results and anaemia symptoms and signs, and achieved higher sensitivity and specificity than the HCS alone. Future studies should therefore include an assessment of combinations of parameters in an attempt to optimize anaemia diagnosis in resource-poor settings.

Although evidence is now available about the accuracy of the HCS, we lack studies addressing issues relevant to widescale implementation and it is therefore difficult to draw conclusions about its usefulness in practice. Future studies need to be pragmatic in design (i.e. to assess diagnostic accuracy in local health care settings with local workers) and should report clearly their method of sampling and characteristics of the patient population. Training methods should be formally evaluated, and should also be practical. Some studies provided up to 2 days training, which is burdensome, expensive, and may not be feasible in practice. The introduction of HCS into settings where there is no laboratory will require careful planning as systems for supplying and safely disposing of sterile lancets will need to be established. Realistic studies in primary care settings are needed which compare the costs and effectiveness of the HCS with the current routine method of clinical diagnosis of anaemia. These studies need to be planned early in the development process and should evaluate longer-term clinical outcomes (such as better management of anaemia, less anaemia at delivery among pregnant women, and increased birth weight), rather than simply reporting on diagnostic accuracy. Such studies will require a multi-disciplinary approach, involving, for example, health planners, social scientists, economists, and epidemiologists as well as clinicians.

### Conclusion

The HCS has the potential to be the most appropriate tool currently available for the detection of mild and moderate anaemias, which are likely to be missed by clinical diagnosis alone, but which still require treatment. It is inexpensive, rapid and simple to use, and the few studies available suggest that under ideal conditions it is reasonably accurate. Although studies are limited, results are promising and justify investment in evaluating its use on a larger scale and in real-life situations. The next step on the pathway to implementation should be to assess its usefulness in combination with clinical diagnosis in practical situations, such as paediatric and ante-natal clinics in developing countries where there is no laboratory. This type of research requires a different approach and research skill-mix from efficacy studies.

### Acknowledgements

This article is based on a background paper commissioned by the World Health Organization for an international expert consultation in May 2004. Funding was also received from the Department for International Development UK (Effective Health Care Research Consortium and Malaria Knowledge Programme). Nevertheless, the views expressed in this article are those of the authors alone. These organizations accept no responsibility for the information or views expressed in this paper. The authors thank Nynke Van Den Broek for supplying additional information from her study. Role of funding source: the funders played no role in study design, collection, analysis, interpretation of data, writing of the report, or in the decision to submit the paper for publication. They accept no responsibility for the contents. Contributions of authors: J.C. assessed studies for inclusion, extracted data, and assessed study methodological quality, carried out analyses, and is responsible for preparing the article. I.B. assessed studies for inclusion, extracted data, and assessed study methodological quality, contributed to analyses, obtained initial funding and helped write the article. Both J.C., and I.B. contributed to the interpretation and revision of the manuscript.

### Conflicts of interest

The authors have declared no conflicts of interest.
Supplementary tables can be found at IJE Online (http://ije.oxfordjournals.org).

KEY MESSAGES
- The HCS is simple, rapid, inexpensive, and suitable for assessing haemoglobin levels on finger prick blood samples in resource-poor settings where there is no laboratory.
- The accuracy of the scale is high in laboratory-based studies but deteriorates as studies become more field-based and ‘real life’, and where prevalence of anaemia is low.
- The HCS is a promising tool for diagnosing anaemia where there is no laboratory, but more research is needed to determine the usefulness of the HCS in ‘real-life’ situations and to assess its effectiveness in improving clinical outcomes.

References

