INCAP Publication I-1107
NUTRITION REPORTS INTERNATIONAL

THE VALUE OF SIMPLE CONJUNCTIVAL EXAMINATION IN FIELD SCREENING FOR ANEMIA

R. Glass, R. Batres, C. Selle, R. Garcia-Ibanez, N. Solomons and F. Viteri

Bureau of Epidemiology, Center for Disease Control, Atlanta Georgia, 30333, U.S.A. and the Division of Human Nutrition and Biology, Institute of Nutrition of Central America and Panama, Guatemala City, Zone 11, Guatemala, Central America

ABSTRACT

Five medically-trained observers with varying amounts of clinical experience participated in a study to assess quantitatively the value of simple conjunctival examination as a field screening procedure for anemia in a population of 85 agricultural workers from the coastal lowlands of Guatemala. On two occasions, one week apart, a numerical estimate of hematocrit was assigned to each subject independently by each observer on the basis of his clinical examination of conjunctival color in full daylight. No specific training, knowledge of the prevalence of anemia in the community, agreement on the hematocrit considered to represent anemia nor information on the true hematocrit of the subjects was provided to the examiners prior to the study. Observers tended to underestimate the true hematocrit values as 89% of laboratory values exceeded the estimates made on the basis of conjunctival color. However, calculating the sensitivity (ability to minimize false negatives) and the specificity (ability to minimize false positives) from a medical decision-making model, all participants had a point of intersection at which both sensitivity and specificity were greater than 70%. Inter-observer variability was less than intraobserver variability and clinical experience had no effect on reliability. The results were compared to those of a similar clinical anemia screening test evaluated in India. It was concluded that individual estimates of hematocrit in unstandardized observers were highly unreliable and that the estimate from a panel of observers is more accurate than repeat observations from the same examiner. However, with feedback training and standardization of observers, the potential for a useful field screening procedure with conjunctival examination is a reasonable expectation.

INTRODUCTION

There are difficulties in estimating the prevalence of anemia in populations, especially in remote areas of developing countries where

expense, technical complexity, or cultural acceptability may place major constraints on the use of conventional indices such as hematocrit or hemoglobin concentration. The reality of this problem has been emphasized in a recent report from India (1). Clinical examination of various membrane surfaces including the tongue, lips, conjunctivae, and nailbeds is a time-honored procedure for screening for anemia at the bedside, but, until recently, neither its application to field situations nor a quantitative evaluation of its utility had been reported. Recently, Ghosh and Mohan (1) evaluated the accuracy of four Indian village health workers in screening for anemia using pallor of lips, tongue, and nailbed. In the present study, we report our assessment of the use of the conjunctival examination in the field as a screening procedure for anemia. We have attempted to determine: 1) the sensitivity and specificity (2) of conjunctival color as a screening test for anemia; 2) its effectiveness for detecting new cases of anemia in a field survey situation; 3) intra- and interobserver variability; and 4) the importance of prior clinical experience as a determinant of reliability.

METHODS

Subjects. The study was conducted in conjunction with a larger hematological assessment and evaluation of iron absorption on a coffee plantation in the lowlands of southern Guatemala (3). The subjects included 85 male agricultural workers. The study was conducted in accordance with the Declaration of Helsinki and the subjects gave their collective consent to be studied after the nature and purpose of the procedures were explained to the assembly. Individuals were excluded from analysis if they had eye disease resulting in conjunctival inflammation or if they failed to appear both for the two examinations and the blood test. A total of 64 men met these criteria and were included in the final tabulations.

Conjunctival examinations. Five medically-trained observers with varying degrees of clinical experience, defined as years of direct patient care either in medical school or in postgraduate practice, participated in the conjunctival examinations. The lower fornix of the conjunctiva was examined in full daylight between 10 and 11 a.m. on both occasions. Instead of the usual binary designations of "anemic" or "not anemic," the observers agreed to give a numerical estimate of the hematocrit of each individual based on the appearance of the conjunctival membrane. Estimates were made and recorded independently without discussion or consultation among observers. At the time of the study, neither the prevalence of anemia nor the hematocrit level which defined anemia for the population were known to the team of examiners.

Hematocrit determination. Venous blood samples were obtained on the first of the two occasions, placed in heparinized tubes, and maintained at 4°C until processing in the laboratory. Microhematocrits were determined with capillary tubes in an International Micro-capillary Centrifuge Model M.B. (International Equipment Company, Boston, Massachusetts) for 6 minutes at 10,000 r.p.m. Hemoglobin concentration was also determined in the same samples by the cyanmethemoglobin method

(Hycel, Inc., Houston, Texas).

Data analysis. The definition of sensitivity and specificity were adapted from Mauser and Bahn (2). Thus, the sensitivity of the conjunctival examination at any estimated hematocrit was defined as the percent of patients with true anemia who were correctly identified by our screening examination. The specificity of the procedure was defined as the percent of patients without anemia who were properly considered to be normal.

RESULTS

Accuracy of hematocrit estimate. The mean hematocrit for the 64 men appearing for both examinations was 41% (range 21 - 55%). The result of a survey in Central America suggested that a hemoglobin concentration of 12 g/dl might be regarded as the lower limit of normal (4); concentrations of 11.9 g/dl or below were regarded as evidence of anemia. A regression curve between hematocrits and hemoglobin of our subjects allowed us to equate a hemoglobin concentration below 11.9 with a hematocrit below 36%. Fourteen of the 64 individuals (22%) had hematocrits below 36%, and were thus considered to be truly anemic. A scattergram of the mean hematocrit for the two separate examinations of the 64 subjects was plotted for each observer. Figure 1 shows a representative scattergram for the estimates for our Observer D. As compared to the true hematocrit, this observer (and all four other participants) consistently tended to estimate values below the theoretical lines of identity for actual versus estimated hematocrit values. Eighty-nine percent of all estimates by all 5 observers underestimated the laboratory values; where overestimation occurred, the subject in question was invariably anemic.

Sensitivity and specificity. For each observer, the sensitivity and specificity, as defined above, were calculated from the scattergram for the full range of estimates (Table 1). Figure 2 illustrates the sensitivity and specificity plots for Observer D whose hematocrit estimates were shown in Figure 1. For example, with regard to specificity, if Observer D had referred all patients whom he estimated to have hematocrits below 30% based on his conjunctival examination for further evaluation, he would have correctly identified 93% of all truly anemic patients in the population and would have left 7% undiagnosed (false negatives). At the same time, his demonstrated specificity would have allowed the correct identification of 60% of non-anemic individuals while the other 40% would have incorrectly been considered as anemic (false positives). Other screening points could have been chosen if one wanted to minimize the number of false positives and maximize the combination of sensitivity and specificity (5). Here we have defined the point of maximal sensitivity and specificity as the intersection point (IP), and use it as a basis for comparing the performance of the 5 different observers. The IP varied by 6 hematocrit units among participants, but all observers achived at least 70% senisivity and specificity at the IP and 3 achieved 82 - 83%.

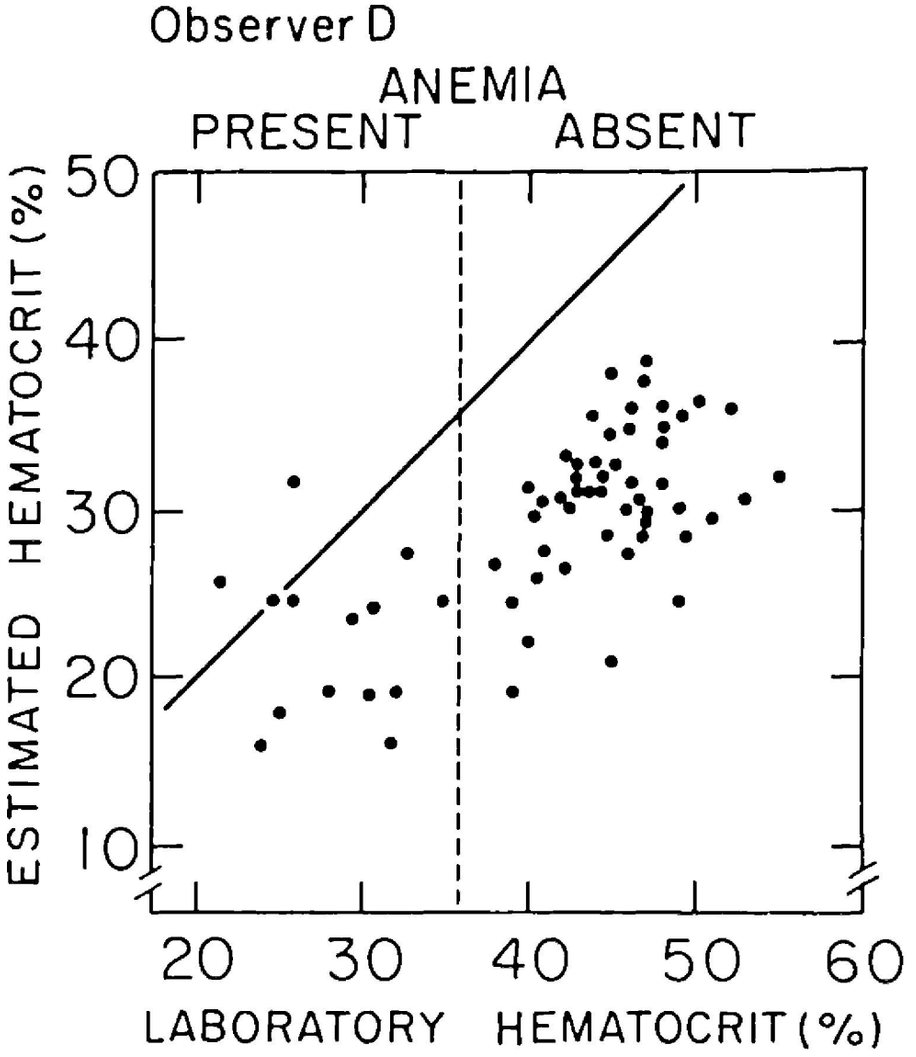

Figure 1. Scattergram of mean hematocrit estimated from conjunctival examination vs. the observed laboratory value for Observer D. The vertical line is the cutoff between "anemica" and "non-anemic" subjects. The diagonal line of identity indicates the curve where the estimated hematocrit is equal to the observed.

TABLE I

COMPARISON OF INTERSECTION POINTS (IP) AT WHICH SENSITIVITY EQUALS SPECIFICITY FOR FIVE OBSERVERS

Observer	Clinical Experience (years)	Hematocrit at IP	Sensitivity & Specificity at IP
A (N.S.)	6	32%	83%
B (R.G.)	4	28%	83%
C (R.G.I.)	2	30%	70%
D (C.S.)	<1	26%	82%
E (R.B.)	<1	31%	76%

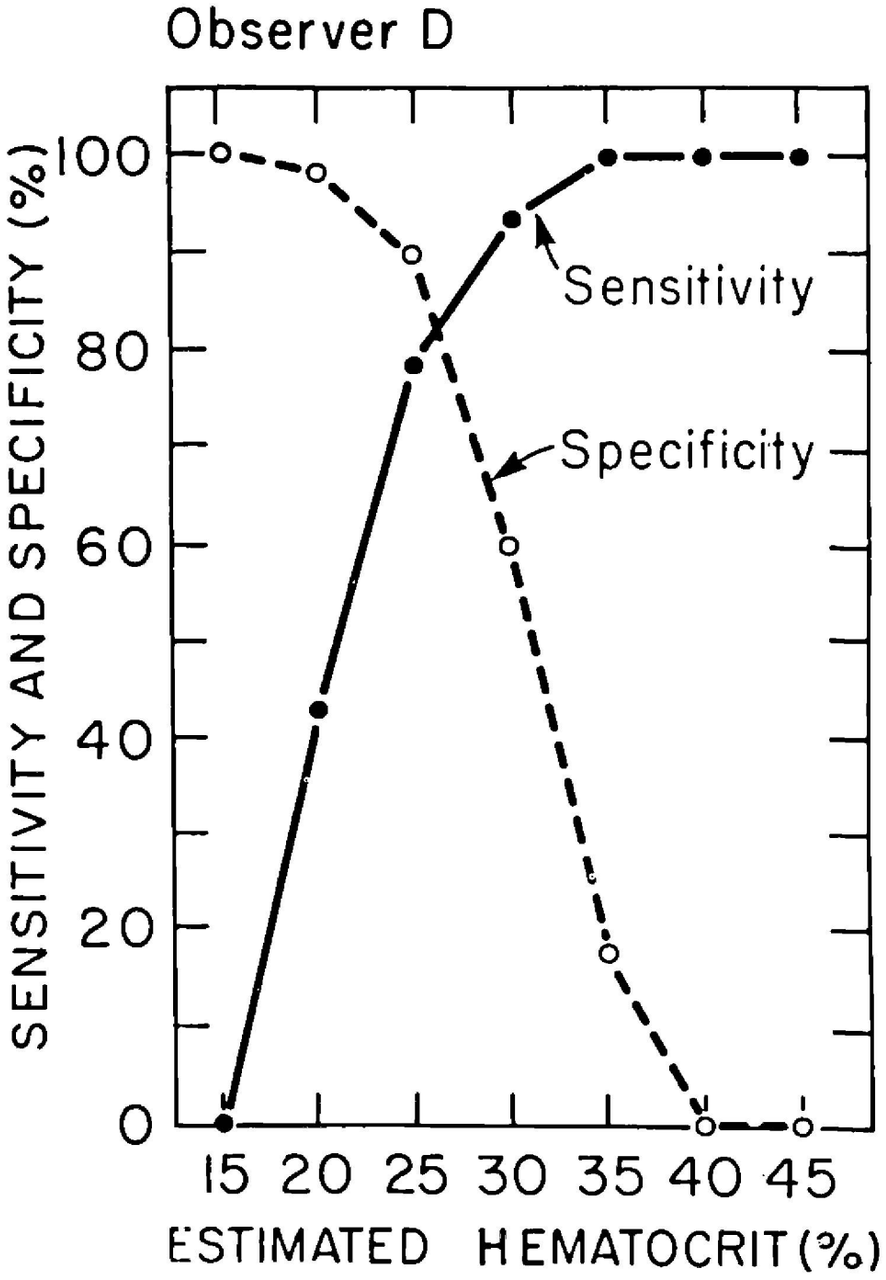

Figure 2. Validity of screening at different estimated hematocrits: Illustration of sensitivity and specificity curves for Observer D.

Intra- and interobserver variability. Five subjects with true hematocrits covering the range for the population were selected to illustrate the variability of estimates from conjunctival examination. The overall intraobserver variability has a standard deviation of 4.0 hematocrit units (about 10% of the average hematocrit reading). The interobserver standard deviation was 2.1 hematocrit units (about 5% of the mean). Thus, the overall panel assessment of hematocrit per patient was more uniform than multiple observations by a single observer.

Influence of clinical experience. The experience of the observer was defined in terms of his years of clinical experience and ranged from 1 for the two medical students to 6 for the board-certified

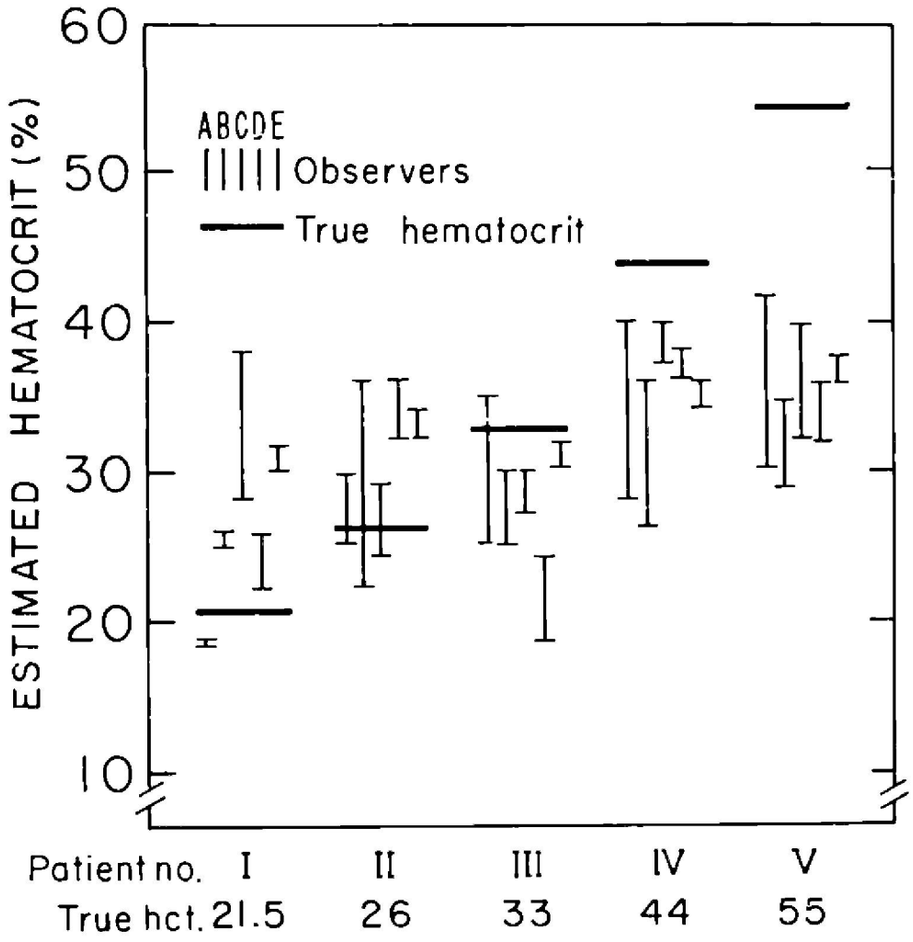

Figure 3. Intra- and interobserver variability: The two hematocrit estimations by each of five observers compared with the true value for five subjects.

internist (see Table 1). The level of experience appeared neither to improve sensitivity and specificity nor to reduce intraobserver variability.

DISCUSSION

With the emphasis today on simplification of the delivery of health care services and the use of locally-trained rural health auxillaries, a quantitative assessment of the clinical evaluation of anemia using the conjunctival examination was considered to be relevant. Strict evaluation of our results illustrated a considerable percentage of our observations were false positive and false negative diagnoses of anemia. We must conclude that, given the conditions of our study, the ability of the observers to estimate accurately the hematocrit of any given individual is highly unreliable with a major tendency to underestimate the actual hematocrit values, and intraobserver variability is large.

Given our experience, just how well would the best observers in the present study have done using the conjunctival examination in a large

field screening project for anemia? The best observer values for sensitivity and specificity, those of Observers A and B, were applied to prevalence data for anemia in Central America reported by Viteri and Guzman (4). In their survey of 1438 adult men living in the lowlands, 63 had hemoglobin values below 11.9 g/dl. Had conjunctival color assessment been used in that survey and had we used the IP as a cutoff point, our best observers would have selected 50 of the 63 subjects with abnormally low hemoglobins as anemic, and each subject screened as positive would have had a 15% chance of truly being anemic (i.e., predictive value (+) = 15%). The greater prevalence of anemia among the population in this study (prevalence = 22%) increased the predictive value of a positive test to 58%. The field trial of clinical anemia evaluation reported by Ghosh and Mohan (1) used pallor of tongue, lower lip, and nailbed as the index. In that study, individuals with hemoglobin concentrations greater than 11 grams per dl were considered not to be anemic. Their four rural health workers correctly identified 100% of subjects with hemoglobin concentrations less than 6 g/dl, and two-thirds of those with hemoglobin in the 6 - 9 g/dl range as anemic. The misdiagnosis of normal individuals as anemic was in the 20 - 30% range.

Despite the objective findings of our present study, i.e. that the hematocrit estimates were fairly unreliable, the data suggest that the potential for developing a more accurate and precise conjunctival evaluation as a field screening procedure for anemia detection is substantial. We chose the cross-over point as the critical value for screening since it maximizes the combination of sensitivity and specificity. We were surprised to find that the validity of conjunctival examination was so high at the point of intersection with all observers achieving cross-over points with greater than 70% sensitivity and specificity. Other cut-off points could reasonably be selected depending upon the tolerance of the investigators for false positives and false negatives in the non-anemic and anemic subpopulations, respectively. In the present study, the 150 conjunctival examinations were made by each observer without knowledge of the true hematocrit. Therefore, there was no feedback mechanism to adjust estimates and improve reliability. If the observer could be standardized to screen individuals at his cross-over point and if the intersection point were adjusted to the cut-off level for anemia in a given population, he could select 70 - 82% of anemic patients for referral for further laboratory confirmation, including those most severely anemic. The Indian study provided color cards as visual references for the health worker. It is quite possible that, given a standard color reference, a training period with immediate feedback, or both, estimates of hematocrits based on conjunctival color might be improved significantly. Theoretically, unlike the color of the lips or nailbeds, which are subject to some variation with an individuals racial origin, occupational history, or hygienic state, or the tongue and lips, which can pick up stains from tobacco smoking, herb chewing, etc., the conjunctivae of individuals without eye disease are relatively clean, unpigmented and physically accessible.

Examination of the conjunctiva is a simple and non-invasive screening test for anemia. Our experience suggests that the observer's ability to distinguish among a large number of individuals with a wide range

of hematocrit values is fair, that interobserver variability is relatively small, and that clinical experience per se does not improve the accuracy of estimation. The ordered numerical estimation of a hematocrit value appears to be more useful than the binary decision between "anemic" and "non-anemic" for the purposes of evaluation of data. Standardization of estimation against actual knowledge of hematocrit values may improve the accuracy and decrease the intra- and interobserver variability for such determinations. If this could be achieved, field examinations of conjunctivae by trained and standardized observers could prove to be an effective method for screening for anemia in populations in which blood sampling is undesirable or unacceptable, and in which anemia is highly prevalent.

ACKNOWLEDGEMENTS

The authors are grateful to Mr. Fabio Saldana, Ms Candelaria Gonzalez, and Ms. Celia Cheth for their field collaboration, to Ms. Cristina de Campos for her laboratory assistance, and to Drs. Harry Smith, T.C. Chalmers, and M. Greenberg for statistical help and critcal review.

REFERENCES

1. Ghosh, S. and Mohan, M. Screening for anemia. Lancet i: 823, 1978.

2. Mausner, J.S. and Bahn, A.K. Epidemiology, an Introductory Text. Saunders, Philadelphia, 1974, p. 237.

3. Viteri, F.E., Garcia-Ibanez, R. and Torun, B. Sodium iron NaFeEDTA as an iron fortification compound in Central America. Absorption studies. Am. J. Clin. Nutr. 31: 961, 1978.

4. Viteri, F.E. and Guzman, M.A. Haematological status of the Central American population. Prevalence of individuals with hemoglobin levels below "normal." Brit. J. Haematol. 23: 725, 1972.

5. McNeil, B.J., Keeler, E. and Adelstein, S.J. Primer on certain elements of medical decision-making. New Engl. J. Med. 293: 211 1975.

Accepted for publication: December 11, 1979.