A Procedure to Estimate the Accuracy and Reliability of HemoCue™ Measurements of Survey Workers

Susan Burger
June Pierre-Louis

A Procedure to Estimate the Accuracy and Reliability of HemoCue™ Measurements of Survey Workers

Susan Burger and June Pierre-Louis

For further information about ILSI's institutes, branches, committees, publications, and research activity, contact:

International Life Sciences Institute
One Thomas Circle, NW, Ninth Floor
Washington, DC 20005-5802 USA
202/659-0074 (phone)
202/659-8654 (fax)
ilsi@ilsi.org (e-mail)
www.ilsi.org

Printed in the United States of America–2003

Preface

The HemoCue™ is widely used to measure hemoglobin in anemia surveys. While an excellent instrument in its own right, data quality is dependent on good blood sample collection—especially where capillary blood is used. One way to ensure that good capillary blood samples are available is to train staff how to stick fingers so that between- and within-person variation in hemoglobin concentrations is minimized. Standardizing field workers is a way to achieve this.

This publication describes a standardization method to estimate the accuracy and reliability of field worker measurements of hemoglobin concentration using the HemoCue™. It is written for managers of programs and surveys that use or intend to use the HemoCue™ to collect hemoglobin measurements to determine the prevalence of anemia. The method described here has been extensively field-tested by Helen Keller International and has been shown to improve the quality of capillary hemoglobin measurements, resulting in more accurate and reliable estimates of anemia prevalence.

Acknowledgments

The Micronutrient Global Leadership project would like to thank Helen Keller International for allowing Drs. Susan Burger and June Pierre-Louis to adapt the material in Chapter V in *How to Assess Iron Deficiency Anemia and Use the HemoCue™* as a stand-alone document. We also acknowledge the help of Dr. Penelope Nestel, Micronutrient Global Leadership, and Ms. Roberta Gutman in this endeavor. The document was reviewed by Dr. Lena Davidsson, Swiss Federal Institute of Technology Zurich, and Mr. Ibrahim Parvanta, U.S. Centers for Disease Control and Prevention.

This publication was made possible with support from Micronutrient Global Leadership, a project of the Office of Health, Infectious Disease and Nutrition, Global Health Bureau, U.S. Agency for International Development, under Cooperative Agreement Number HRN-A-00-98-00027-00 with the International Life Sciences Institute (ILSI) Research Foundation. The original publication from which the present publication was excerpted and adapted, *How to Assess Iron Deficiency Anemia and Use the HemoCue™* was made possible with support from the U.S. Agency for International Development under contract number FAO-0500-A-00-6001-00 with Helen Keller International.

A Procedure to Estimate the Accuracy and Reliability of HemoCue™ Measurements of Survey Workers

Susan Burger and June Pierre-Louis

The HemoCue™ measures hemoglobin (Hb) concentration in whole blood. HemoCue™ measurements taken from individuals are used to determine the extent to which anemia exists in a population. With correct instrument use, the HemoCue™ can give highly accurate results that compare with standard laboratory assays using venous blood (von Schenck et al. 1986, Lardi et al. 1998). HemoCue™ Hb concentration in capillary blood can provide an adequate estimate of population anemia prevalence compared with venous blood, irrespective of analysis method (HemoCue™ compared with, e.g., an automated spectrophotometer), but sample collection technique is critical (Neufeld et al. 2002). Good sample collection and analysis technique ensures that measurements are both accurate and reliable. *Accuracy* is the closeness of an Hb concentration measurement to its true value; *reliability* is the closeness of repeated measurements of Hb concentration.

Good training and practice to minimize intra- (within) and inter- (between) observer variations in measurements when using the HemoCue™ is important for taking accurate and reliable Hb measurements. It is also essential that those taking HemoCue™ measurements meet certain performance standards. If HemoCue™ measurements are not accurate and reliable, the estimates of the extent of anemia in a population will likewise not be correct. To maximize the likelihood of achieving accurate and reliable readings, it is important to follow a protocol to "standardize" procedures when training field workers to measure capillary Hb using the HemoCue™.

Because verified quality control material is not available for the HemoCue™, the standardization procedure described in this document shows how closely the Hb measurements taken by a trainee are to those of an experienced trainer. It also helps pinpoint why errors are made so that they can be eliminated or at least minimized. The standardization procedure also enables a survey manager to pick the most competent survey workers to conduct HemoCue™ measurements in the field. Repeated standardization of survey workers during data collection will also verify that errors are minimized. The procedures in this document were extensively field-tested by private voluntary organizations in Bangladesh, Mali, Mozambique, Niger, and Peru. It is important, however, to note that neither the procedures described here nor the use of the control cuvette can detect systematic errors if both the trainer and the trainee are making the same systematic error. Careful training should minimize the possibility that all of the survey workers and trainers are making the same systematic error.

Sticking fingers and getting a large enough blood drop to fill the microcuvette are two of the most difficult and error-prone steps in using the HemoCue™ and can ultimately result in incorrect estimates of the extent of anemia in a population. Information on how to take capillary blood samples is described elsewhere (Sharmanov 2000, Burger and Pierre-Louis 2002).

Practice Exercise

The standardization procedure is preceded by training and practice, during which blood samples are collected and measured and the skill of the trainee determined. Trainers should also complete this practice exercise to refresh their skills before beginning the training. Trainers and trainees should practice taking two finger sticks from the same person and then practice taking two blood drops from one finger stick.

Step A. Each trainer and trainee uses the procedure described in Table 1 to take a blood drop from two separate finger sticks from each volunteer. The trainer and trainee practice until the samples from these duplicate finger sticks are within 0.5 g/dL on two consecutive volunteers.[1] Generally, five volunteers are sufficient for most people to practice on. A comparison of the results as recorded on Part A of Form 1 indicates how good the individual is at finger sticking. Example 1 shows how the form is used to record data for this comparison.

Step B. Then each trainer and trainee takes two blood drops from a single finger stick from each volunteer until the Hb measurements from duplicate blood drops are within 0.5 g/dL on two consecutive volunteers. Again, five volunteers are generally sufficient for practicing duplicate blood drops. Part B of Form 1 is used for this purpose.

Data Collection

1. The name of the trainer or trainee is recorded at the top of the form.
2. For each volunteer, the trainer and trainee uses the procedure described in Table 1 to collect one blood sample from each of two finger sticks (Form 1, Part A) or two blood samples from one finger stick (Form 1, Part B). The HemoCue™ reading for the first blood sample is recorded in column I of the first row, and the reading for the second sample is recorded in column II of the same row.
3. The HemoCue™ reading of the second sample (column II of Form 1) is then subtracted from the reading of the first sample (column I), and the difference, whether positive (+) or negative (–), is recorded in column III in the same row. This is the difference between the two measurements taken by the trainee.
4. The trainer and trainee circle any differences (positive or negative) that are greater than or equal to 0.5 g/dL and discuss the possible reasons for the differences (see the boxed information on page 4).[3] Suspected reasons for these differences are recorded in column IV.

Table 1. Procedure for collecting one or two blood samples from each finger stick[2]

	Form 1: Part A Form 2: Trainer	Form 1: Part B Form 2: Trainee
Stick the selected finger with the lancet.	✓	✓
Wipe away the first drop of blood with a clean gauze pad.	✓	✓
Wipe away the second drop of blood with the same gauze pad.	✓	✓
Fill the first microcuvette with the third drop.	✓	✓
Quickly wipe away any remaining blood from the third drop.		✓
Fill the second microcuvette with the fourth drop.		✓
Read the first microcuvette.	✓	✓
Read the second microcuvette.		✓

[1] The criterion of 0.5 g/dL was developed in consultation with a group of technical advisers who had conducted training sessions for using the HemoCue™ for population-based surveys.

[2] The present document describes how to use the HemoCue™ in the practice and standardization exercises. A Quality Assurance Logbook and additional checklists for ordering appropriate supplies, maintaining the HemoCue™ to optimize its accuracy, maintaining a clean workstation, protecting survey workers and survey subjects from exposure to blood, and sticking a finger, as well as detailed instructions on how to fill a microcuvette, are available in Burger and Pierre-Louis (2002, chap. 4 ["How Do I Use the HemoCue™ to Measure Hemoglobin?"]).

[3] A full range of common problems that have occurred while filling the microcuvette during field training is described in Burger and Pierre-Louis (2002, chap. 4).

Form 1. Practice to get reliable hemoglobin values using the HemoCue™

Trainer/Trainee: .....

Part A. Duplicate finger sticks: one blood drop from each of two finger sticks from each volunteer

Volunteer	I	II	III	IV
	Sample 1	Sample 2	Sample 1 (column I) – sample 2 (column II)	Potential reasons for differences ≥ 0.5 g/dL
1				
2				
3				
4				
5				

Part B. Duplicate blood drops: two blood drops from one finger stick from each volunteer

Volunteer	I	II	III	IV
	Sample 1	Sample 2	Sample 1 (column I) – sample 2 (column II)	Potential reasons for differences ≥ 0.5 g/dL
6				
7				
8				
9				
10				

Note: Record whether the differences are positive (+) or negative (–) in Column III.

Example 1. Practice to get reliable hemoglobin values using the HemoCue™

Trainer/Trainee: *Mohamed*.....

Part A. Duplicate finger sticks: one blood drop from each of two finger sticks from each volunteer

Volunteer	I	II	III	IV
	Sample 1	Sample 2	Sample 1 (column I) – sample 2 (column II)	Potential reasons for differences ≥ 0.5 g/dL
1	*13.4*	*14.1*	*–0.7*	*Squeezed finger on 1st finger stick*
2	*10.7*	*10.7*	*0.0*	
3	*14.7*	*13.4*	*1.3*	*Air bubbles in 2nd microcuvette*
4	*14.9*	*14.7*	*0.2*	
5	*16.0*	*15.8*	*0.2*	

Part B. Duplicate blood drops: two blood drops from one finger stick from each volunteer

Volunteer	I	II	III	IV
	Sample 1	Sample 2	Sample 1 (column I) – sample 2 (column II)	Potential reasons for differences ≥ 0.5 g/dL
6	*14.8*	*14.9*	*–0.1*	
7	*16.4*	*15.5*	*0.9*	*Squeezed finger for 2nd blood drop*
8	*16.9*	*17.4*	*–0.5*	*Squeezed finger for 1st blood drop*
9	*11.1*	*11.2*	*–0.1*	
10	*12.4*	*12.2*	*0.2*	

Note: Record whether the differences are positive (+) or negative (–) in Column III.

Three kinds of errors commonly cause **false low** Hb concentration readings:

- The finger was still wet from the alcohol solution when it was punctured, which diluted the blood with alcohol.
- The finger was squeezed hard or "milked," which diluted the blood drop with interstitial fluid.
- The microcuvette contained air bubbles (seen when held up to the light), which lowered the concentration of red blood cells in the microcuvette.

Three kinds of errors commonly cause **false high** Hb concentration readings:

- The microcuvette was incompletely filled (seen when held up to the light) because of poor blood flow from a shallow finger stick. Chemicals in the microcuvette thus did not mix properly with the reagents.
- The microcuvette was taken from a container that had been opened more than one week ago in a hot climate (or more than three months ago in a temperate climate), resulting in deterioration of the chemicals in the microcuvette.
- The blood sample clotted before the microcuvette was filled, causing the blood to be more concentrated than it should have been, primarily because it took too long to form an adequate blood drop. That is, the blood drop did not form quickly enough.

Standardization Procedure

Trainers are strongly urged to carry out this standardization procedure with a group of colleagues before training sessions are begun. Carrying out the standardization procedure among trainers can uncover systematic errors even among experienced personnel. For example, during field-testing of this procedure, one experienced trainer discovered she was wiping the finger with alcohol after sticking the finger, which invalidated the findings of an entire survey.

Each trainee takes duplicate blood drops[4] from one finger stick from each of 10 volunteers.[5] The trainer then takes a sample from a separate finger stick from each of the volunteers. A comparison of the results as recorded on Form 2 shows the within- and between-observer variation in measurements. The data from this exercise are used to conduct the standardization procedure that will provide quantitative measures to help survey managers choose skilled and reliable trainees to serve as survey workers. The data in Form 2 are for the trainees to review with the trainer, and are collected and analyzed as follows:

Data Collection

1. The names of the trainee and trainer are recorded at the top of Form 2.
2. For each volunteer, the trainee uses the procedure described in Table 1 to collect two blood samples from a single finger stick. The HemoCue™ reading for the first blood sample is recorded in column I of row 1, and the reading for the second sample is recorded in column II in the same row.
3. From the same volunteer, the trainer collects a

[4] Strictly speaking, accuracy and reliability should be measured from four finger sticks on each volunteer. Because it is impractical to take four finger sticks from one person, reliability is assessed using duplicate blood drops, for two reasons. First, finger sticks cause pain and discomfort and, second, once finger sticking is mastered, it is possible to fill two microcuvettes with two consecutive blood drops from one finger stick. If the finger stick is not deep enough, the blood flow will be inadequate to fill two microcuvettes

[5] The volunteers should reflect the age and gender of the survey subjects. If children are to be surveyed, trainees should also practice on children. A discussion of the selection of appropriate target groups for a survey is found in Burger and Pierre-Louis (2002, chap 3 ["How Do I Conduct an Anemia Survey?"].

separate sample from another finger stick. This third HemoCue™ reading is recorded in column V of row 1.

4. The HemoCue™ reading of the second sample (column II) is then subtracted from the reading of the first sample (column I), and the difference, whether positive (+) or negative (–), is recorded in column III in the same row. This is the difference between the two measurements taken by the trainee.
5. The HemoCue™ reading for the third sample (column V) is then subtracted from the first sample (column I), and the difference, whether positive or negative, is recorded in column VI of the same row. This is the difference between the first measurement taken by the trainee and the measurement taken by the trainer. Steps 2 through 5 are repeated for the remaining volunteers.
6. All values in columns III and VI ≥ 0.5 g/dL, whether positive or negative, are circled. The trainer and trainee should review every circled value and discuss the potential reasons why the HemoCue™ readings could have been erroneously high or low. The trainer should record the most plausible reasons in columns IV and VII. The trainer and trainee should also look at whether there is a consistent pattern of negative or positive errors, which may lead to the discovery that the trainee is repeating the same mistake. A mix of positive and negative errors can mean that the trainee is making different mistakes at different times.
7. The sample means are then calculated as follows:
   a. The readings for the first sample taken from each of the volunteers (the values in column I of Form 2) are added and the sum recorded at the bottom of the column in the row marked sum. The same is done for the second and third samples (the values in column II and the values in column V).
   b. Each of the three sums is divided by the number of volunteers (generally 10), and the result is recorded at the bottom of columns I, II, and V in the row marked mean. The difference between any of the

Form 2. Standardization exercise to get accurate and reliable hemoglobin values using the HemoCue™

Trainee: ..... **Two blood drops from a single finger stick**

Trainer: ..... **A third blood drop from a second finger stick**

	I	II	III	IV	V	VI	VII
Volunteer	**Trainee, blood drop 1**	**Trainee, blood drop 2**	**Column I – column II**	**Potential reasons for differences ≥ 0.5 g/dL**	**Trainer, blood drop 3**	**Column I – column V**	**Potential reasons for differences ≥ 0.5 g/dL**
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
Sum							
Mean							

Note: Record whether the differences are positive (+) or negative (–) in columns III and VI.

Example 2. Standardization exercise to get accurate and reliable hemoglobin values using the HemoCue™

Trainee: *Fatima* ..... **Two blood drops from a single finger stick**

Trainer: *Abu Baker*..... **A third blood drop from a second finger stick**

Volunteer	I	II	III	IV	V	VI	VII
	Trainee, blood drop 1	Trainee, blood drop 2	Column I – column II	Potential reasons for differences ≥ 0.5 g/dL	Trainer, blood drop 3	Column I – column V	Potential reasons for differences ≥ 0.5 g/dL
1	9.4	9.7	–0.3		9.7	–0.3	
2	11.0	11.6	(–0.6)	2nd microcuvette not completely filled	11.3	–0.3	
3	12.0	12.4	–0.4		12.8	(–0.8)	Trainee squeezed finger while filling microcuvette
4	10.9	11.5	(–0.6)	2nd microcuvette not completely filled	11.6	(–0.7)	Trainee squeezed finger while filling microcuvette
5	12.6	12.5	0.1		11.8	(0.8)	Air bubbles in trainer's microcuvette
6	13.5	14.2	(–0.7)	2nd microcuvette not completely filled	13.4	0.1	
7	13.7	13.4	0.3		13.3	0.4	
8	10.0	12.8	(–2.8)	Alcohol not dry before filling 1st microcuvette	13.3	(–3.3)	Alcohol not dry before trainee filled microcuvette
9	10.0	10.5	(–0.5)	2nd microcuvette not completely filled	10.2	–0.2	
10	13.9	13.3	(0.6)	1st and 2nd microcuvette not completely filled	10.2	(3.7)	Trainee's microcuvette not completely filled
Sum	117.0	121.9			117.6		
Mean	11.7	12.2			11.8		

Note: Record whether the differences are positive (+) or negative (–) in columns III and VI.

means should be no greater than 0.5 g/dL. If the difference is ≥ 0.5 g/dL, the trainee needs further practice to improve his or her sampling technique before being considered as a field worker. After the trainee receives further training, the standardization procedure will need to be done again to determine whether he or she can be considered for field work.

Data Evaluation

Large differences between the HemoCue™ readings of duplicate blood samples are an indication of *unreliability*. In Example 2 (a filled-out Form 2), the differences between the measurements of the trainee's duplicate samples from six of the 10 volunteers were ≥ 0.5 g/dL, which indicates that this trainee did not measure Hb concentrations reliably. Large differences that are consistently positive or negative may indicate that the trainee has not yet mastered the process of collecting blood samples for the HemoCue™. Because there were more negative (five) than positive (one) differences, the type of errors made by the trainee (not filling the microcuvette and squeezing the finger) might have been the result of shallow finger sticks.

Large differences between the blood samples measured by the trainee and those measured by the trainer that are consistently positive or negative indicate unacceptable interobserver variation in blood collection technique and are *inaccurate*.[6] In Example 2, five of the 10 differences between the

[6] Ideally, accurate measurements are made using venous blood and a "gold standard" laboratory test for anemia. Both are often impractical to implement in field settings, so capillary measurements taken by an experienced trainer who has already undergone the practice and standardization and who, for practical reasons, is assumed to do good finger pricks are used as a more feasible alternative.

measurements taken by the trainee and the trainer were ≥ 0.5 g/dL, and most of these (three out of five) were negative. The trainer made only one error (on volunteer 5) whereas the trainee made four errors, which indicates that this trainee's measurements were inaccurate—using these adapted procedures—as well as unreliable.

To select trainees to function as survey workers capable of taking reliable measurements, the trainer must evaluate the capability of the trainees based on the reliability and accuracy of their measurements using the adapted procedure to optimize HemoCue™ measurements. **Here, *reliability* is estimated by calculating the *intra*observer variance and *accuracy* by calculating the *inter*observer variance.**

Form 3 is used by the trainer to calculate the intraobserver variance (the variance of the differences between the duplicate samples measured by the trainee) and the interobserver variance (the variance of the differences between the samples measured by the trainee—in this exercise, the first samples—and those measured by the trainer).[7] Example 3 (a filled-out Form 3) illustrates how variances of HemoCue™ measurements are determined.

8. Calculate the sum of the squares by first squaring each number in column III of Form 3 (e.g., $0.3 \times 0.3 = 0.09$); record these numbers in column IV. Add the numbers in column IV and record the sum at the bottom of column IV in the row marked sum. This is the first sum of the squares. Then square each number in column

Form 3. Variances of the HemoCue™ measurements

Trainee: ..... **Two blood drops from a single finger stick**

Trainer: ..... **A third blood drop from a second finger stick**

Volunteer	I Trainee, blood drop 1	II Trainee, blood drop 2	III Column I – column II	IV (Column III)^2	V Trainer, blood drop 3	VI Column I – column V	VII (Column VI)^2
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
Sum							
Mean							
Variance			Intraobserver			Interobserver	

Note: Record whether the differences are positive (+) or negative (–) in columns III and VI.

[7] Ideally, intra- and interobserver errors would be better indicators for standardizing a trainee's measurements. To calculate intra- and interobserver error, each volunteer would need to have his or her fingers stuck four times: two finger sticks and samples collected by the trainee and two finger sticks collected by the trainer. Four finger sticks are not well tolerated by volunteers. This adapted procedure is a feasible alternative that works well in field conditions.

Example 3. Variances of the HemoCue™ measurements

Trainee: *Fatima* ..... **Two blood drops from a single finger stick**

Trainer: *Abu Baker* ..... **A third blood drop from a second finger stick**

Volunteer	I Trainee, blood drop 1	II Trainee, blood drop 2	III Column I – column II	IV (Column III)^2	V Trainer, blood drop 3	VI Column I – column V	VII (Column VI)^2
1	*9.4*	*9.7*	*–0.3*	*0.09*	*9.7*	*–0.3*	*0.09*
2	*11.0*	*11.6*	*–0.6*	*0.36*	*11.3*	*–0.3*	*0.09*
3	*12.0*	*12.4*	*–0.4*	*0.16*	*12.8*	*–0.8*	*0.64*
4	*10.9*	*11.5*	*–0.6*	*0.36*	*11.6*	*–0.7*	*0.49*
5	*12.6*	*12.5*	*0.1*	*0.01*	*11.8*	*0.8*	*0.64*
6	*13.5*	*14.2*	*–0.7*	*0.49*	*13.4*	*0.1*	*0.01*
7	*13.7*	*13.4*	*0.3*	*0.09*	*13.3*	*0.4*	*0.16*
8	*10.0*	*12.8*	*–2.8*	*7.84*	*13.3*	*–3.3*	*10.89*
9	*10.0*	*10.5*	*–0.5*	*0.25*	*10.2*	*–0.2*	*0.04*
10	*13.9*	*13.3*	*0.6*	*0.36*	*10.2*	*3.7*	*13.69*
Sum	*117.0*	*121.9*		*10.01*	*117.6*		*26.74*
Mean	*11.7*	*12.2*			*11.8*		
Variance			**Intraobserver**	*0.50*		**Interobserver**	*1.34*

Note: Record whether the differences are positive (+) or negative (–) in columns III and VI.

VI and record these numbers in column VII. Add the numbers in column VII and record the sum at the bottom of column VII in the row marked sum. This is the second sum of the squares.

9. Calculate the variance by dividing the sums at the bottom of columns IV and VII by twice the number of volunteers (in this case 10 volunteers; i.e., divide by 20 [10×2]). Record these values in the row marked variance at the bottom or columns IV and VII. If these values are reliable, the intraobserver variance in column IV should be less than 0.5 and the interobserver variance in column VII should be less than 1.0. (Extensive field experience has shown that these are appropriate cutoff figures to use in selecting competent field workers who will take accurate and reliable measurements.) Note that in

Table 2. Variances of HemoCue™ measurements

Trainee	Volunteer(n)	Variance Intraobserver (trainee vs. self)	Variance Interobserver (trainee vs. trainer)
A	10	0.09	0.71
B	10	0.23	0.36
C	12	0.11	0.29
D	10	0.33	0.96
E	10	1.80	1.69
F	10	0.32	0.21
G	10	0.17	0.67
H	11	1.32	1.17

Example 3, the interobserver variance is 1.34 and the intraobserver variance is 0.50, so the measurements performed by this trainee are unreliable and inaccurate.

Table 2 shows variances from an actual training exercise. The intraobserver variance of duplicate samples measured by the trainees was usually less than the interobserver variance of the samples measured by the trainee and those measured by the trainer. Individuals E and H did not have variances low enough to meet the standardization criteria (<0.5 for intraobserver variance and <1.0 for interobserver variance). The variances of these trainees indicated that their measurements were neither reliable nor accurate. These two trainees thus were required to continue practicing until they met the criteria for a survey worker.

The Effect of Measurement Errors on Estimates of Anemia Prevalence

When survey workers measure the Hb concentration of individual subjects with greater reliability and accuracy, the estimate of the prevalence of anemia will be closer to the true prevalence. When measurement errors are unreliable, the estimated prevalence of anemia is likely to be higher than the actual prevalence. When measurement errors are consistently inaccurate, the estimated prevalence of anemia may be either higher or lower than the actual prevalence.

Errors in technique that produce readings on the HemoCue™ that are lower than the true Hb concentration lead to *overestimates* of the prevalence of anemia. Errors in technique that produce readings on the HemoCue™ that are higher than the true Hb concentration lead to *underestimates* of the prevalence of anemia.

Measurement error can be due to unreliability or inaccuracy or both. Figure 1 illustrates reliable or consistent Hb concentration measurements from a training exercise using the procedures described here. A trainee collected and measured two blood samples from each of five survey subjects. The solid circles represent the values for sample 1, and the open circles represent the values for sample 2. The difference between the two measurements from each of the same subjects is less than 0.5 g/dL. The solid line represents the mean of the first five samples (12.6 g/dL), and the dotted line represents the mean of the second five samples (12.8 g/dL). Because the mean values for the first and second samples are very similar, this trainee's measurements of Hb concentration were considered reliable.

Figure 1. Reliable measurements of Hb concentration using the HemoCue™

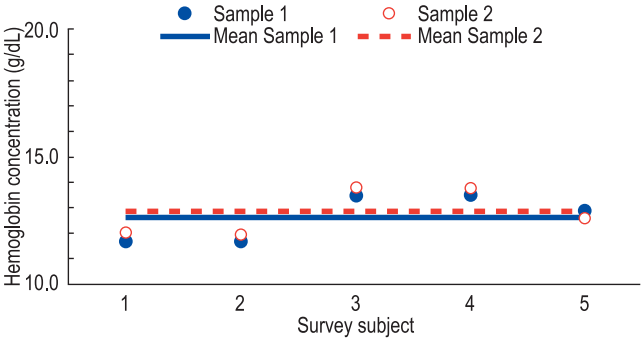


Figure 2. Unreliable measurements of Hb concentration using the HemoCue™

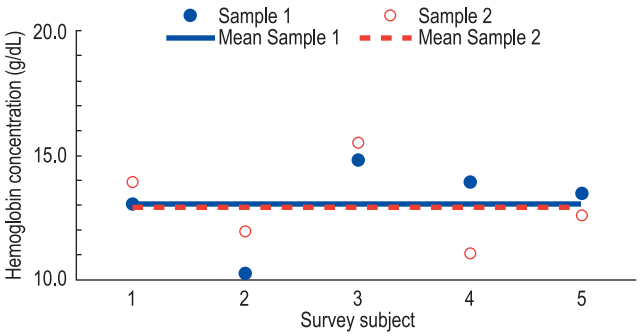


Figure 2 illustrates unreliable or inconsistent measurements. As before, a trainee collected and measured two samples from each of five survey subjects. Here, the difference between the two measurements from each of the subjects is greater than 0.5 g/dL. The mean values of the first and second sample are, however, similar at 13.1 and 13.0 g/dL, respectively.

Because the mean values of the unreliable measurements in Figure 2 are similar, it might be questioned whether reliable measurements are important. The answer is yes, they are, for the reasons given below.

Figure 3 shows what can happen if the same population was measured by two different survey

pairs or teams, one that took reliable measurements and one that did not. The solid triangles show the prevalence estimate of team 1, which took reliable Hb measurements. The open triangles show the prevalence estimate of team 2, which took unreliable Hb measurements. The World Health Organization (WHO) cutoff value for anemia in children 6–59 months is also shown (11 g/dL).

The estimated mean Hb values for team 1 (11.9 g/dL) and for team 2 (11.7 g/dL) are similar, but the distribution curves of the measurements taken by each team are quite different. The distribution curve of team 1 is taller and has relatively fewer values at the tails of the curve, reflected in flatter tails. The distribution curve of team 2 is much shorter with higher and longer tails. This is because some of the unreliable measurements are smaller than the true values whereas others are larger than the true values. The smaller and larger values will make the curve wider. Thus, the estimated prevalence of anemia according to team 2 (the open triangles, representing unreliable measurements) is 33% because the left-hand tail of this distribution curve is longer, whereas the estimated prevalence of anemia according to team 1 (the solid triangles, representing reliable measurements) is only 21%.

Measurements must be accurate as well as reliable. Accuracy is how close a measurement of Hb concentration is to the true Hb concentration. Based on the procedures described here, Figure 4 illustrates inaccurate measurements taken by a trainee who collected and measured two blood samples from each of five survey subjects. The circles represent the values for the trainee, and the squares represent the values for the trainer. None of the two Hb measurements from the same subjects are within 0.5 g/dL of each other. The dashed line represents the mean of the five samples taken by the trainee (12.6 g/dL), and the solid line represents the mean of the five samples taken by the trainer (11.7 g/dL). The mean of the trainee is more than 0.5 g/dL greater than the mean of the trainer.

Figure 5 shows what could happen if the same population was measured by two different survey teams, one that took accurate measurements and one that did not. The solid triangles show the prevalence estimate of team 1, which took accurate

Figure 3. Distribution of Hb concentrations in preschool children (6–59 months) according to the reliability of HemoCue™ measurements

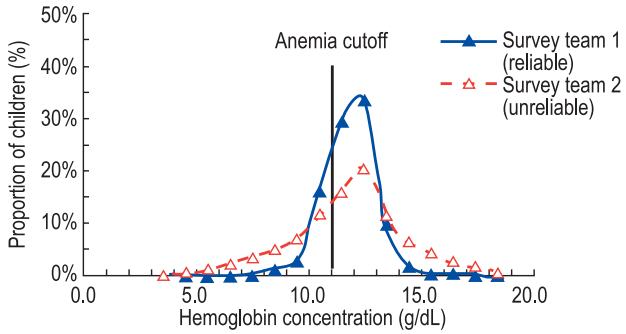


Figure 4. Inaccurate measurements of Hb concentration using the HemoCue™

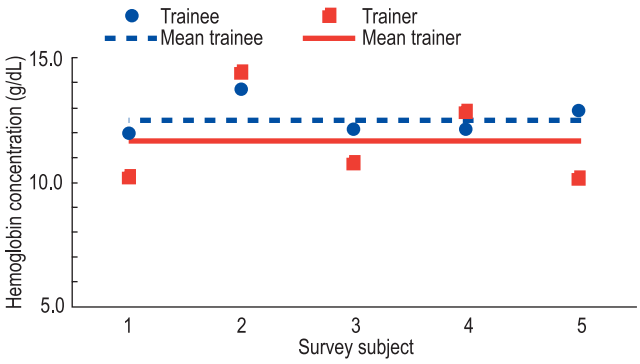
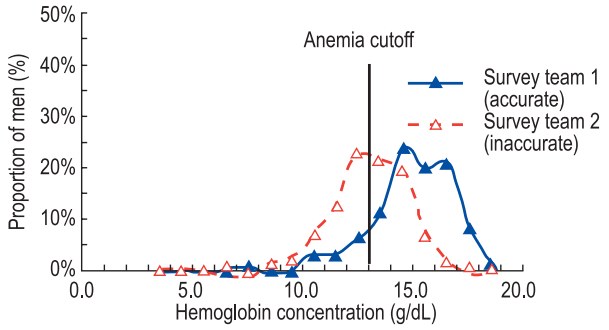


Figure 5. Distribution of Hb concentrations in adult men according to the accuracy of HemoCue™ measurements



Hb measurements. The open triangles show the prevalence estimate of team 2, which took inaccurate Hb measurements. The WHO cutoff value for anemia in adult men is also shown (13 g/dL).

The shape of the distribution curve of team 1 is similar to the shape of the curve of team 2. However, the distribution curve of team 2 is shifted to the left of the curve of team 1. Measurement error usually leads to overestimates of anemia prevalence. The mean Hb concentration of team 1 is 14.9 g/dL,

which is quite different from the mean Hb concentration of team 2, 13.9 g/dL. The difference in the estimated prevalence is very large because the distribution curve for the inaccurate measurements is to the left of the curve for the accurate measurements. The estimated anemia prevalence is 49% based on the inaccurate measurements but is only 14% based on the accurate measurements.

Maintaining Measurement Quality Throughout a Survey

Ideally, survey workers work in pairs, one taking measurements while the other records data and prepares the supplies. Each survey worker should be able to do both activities. In addition to thorough training, the quality of the measurements taken by the survey workers must be maintained *throughout* the survey using the standardization procedure described above.

One way the quality of the Hb measurements can be maintained throughout the survey is by systematically selecting a small proportion of the survey subjects from whom two finger sticks will be taken by each of the two survey workers. In a survey, typically 15 subjects are selected by a random-start systematic method in 30 clusters for a total of 450 subjects for each age and sex group in each target area. A subsample of one out of 15 survey subjects is selected by a random process to undergo two finger sticks for the standardization procedure. If the chosen subject does not consent, the next subject can be selected. The intra- and interobserver variance is calculated once data are available for a subsample of 10 subjects (out of 150 total subjects). Ideally, data quality is verified three times during the survey from a total subsample of 30 subjects. In practice, however, many survey subjects may not find it acceptable to have two finger sticks taken.

Another way to maintain quality is by randomly choosing one survey worker in each pair to collect a blood sample from survey participants 1, 3, 5, and so on, while the other survey worker collects a sample from participants 2, 4, 6, and so on, until the survey on that particular day is completed. Each survey worker in a pair thus measures half of the samples on a particular day. The mean Hb values of the samples taken by the two survey workers are then compared. They should be very close to each other if the survey workers are taking reliable measurements and if the sampling of subjects is representative of the target age/sex group in that community.[8] As mentioned above, typically 15 subjects are selected by a random-start systematic method in 30 clusters for a total of 450 subjects in each age and sex group and in each target area to be surveyed. Using this method for monitoring survey worker reliability, the Hb distribution curves are checked three times throughout the survey, when data have been collected from a total of 150 subjects (75 each). The Hb distribution curves will be very similar if the survey workers are taking reliable measurements and the sampling of subjects is representative of the target age/sex group in that community. If the distribution curve of the Hb values for one survey worker is much wider and flatter than that of the other survey worker, the measurements of the survey worker who produced the wider and flatter distribution curve are not reliable. This worker thus must be retrained and restandardized, or replaced.

References

Burger S, Pierre-Louis J. How to assess iron deficiency anemia and use the HemoCue™. New York: Helen Keller International, 2002. Available at: http://www.hki.org/publications/index.html. Accessed December 1, 2003.

Lardi AM, Hirst C, Mortimer AJ, McCollum CN. Evaluation of the HemoCue for measuring intra-operative haemoglobin concentrations: a comparison with the Coulter Max-M. Anaesthesia 1998;53:349–352.

Neufeld L, Garcia-Guerra A, Sanchez-Francia D, Newton-Sanchez O, Ramirez-Villalobos MD, Rivera-Dommarco J. Hemoglobin measured by HemoCue and a reference method in venous and capillary blood: a validation study. Salud Publica Mex 2002;44:219–227.

Sharmanov A. Anemia testing manual for population-based surveys. Calverton, MD: Macro International Inc., 2000.

Von Schenck H, Falkensson M, Lundberg B. Evaluation of "HemoCue," a new device for determining hemoglobin. Clin Chem 1986;32:526–529.

[8] A discussion of the practical issues of sampling for population-based surveys of anemia can be found in Burger and Pierre-Louis (2002, chap. 3).

This publication is made possible by support from Micronutrient Global Leadership, a project of the Office of Health, Infectious Disease and Nutrition, Global Health Bureau, U.S. Agency for International Development, under Cooperative Agreement Number HRN-A-00-98-00027-00.

Additional copies of this and other Micronutrient Global Leadership publications are available free of charge to developing countries and for US$3.50 to developed countries. Copies can be ordered from Micronutrient Global Leadership at:

Micronutrient Global Leadership
ILSI Human Nutrition Institute
One Thomas Circle, NW
Ninth Floor
Washington, DC 20005, USA

Tel: (202) 659-9024
Fax: (202) 659-3617
Email: hni@ilsi.org
Internet: http://hni.ilsi.org