The purpose of the International Nutritional Anemia Consultative Group (INACG) is to guide international activities aimed at reducing nutritional anemia in the world. The group offers consultation and guidance to various operating and donor agencies who are seeking to reduce iron deficiency and other nutritionally preventable anemias. As part of this service, INACG has prepared guidelines and recommendations for:

—assessing the regional distribution and magnitude of nutritional anemia,

—developing intervention strategies and methodologies to combat iron deficiency anemia,

—evaluating effectiveness of intervention programs on a continuing basis so that evaluation is a continuing and dynamic procedure;

—research needed to support the assessment, intervention and evaluation of programs.

Monographs published by the International Nutritional Anemia Consultative Group are

—Guidelines for the Eradication of Iron Deficiency Anemia
—Iron Deficiency in Infancy and Childhood
—Iron Deficiency in Women
—The Effects of Cereals and Legumes on Iron Availability

These reports are available from:

The Nutrition Foundation, Inc.
888 Seventeenth Street, N.W.
Washington, D.C. 20006
U.S.A.

ISBN 0-935368-44-2
October 1984

PNAAS-425
39298

THE DESIGN AND ANALYSIS OF IRON SUPPLEMENTATION TRIALS

A Report of the International Nutritional
Anemia Consultative Group

AUTHORS

Dr. Selwyn J. Baker
Professor of Medicine
University of Manitoba
St. Boniface General Hospital
Winnipeg, Manitoba, R2H 2A6
CANADA

Mr. K. Ramachandran
Associate Professor and Head,
Biostatistics Division
All India Institute of
Medical Sciences
Ansari Nagar, New Delhi 29
INDIA

Officers of INACG Editorial Review Board

Dr. Edouard M. DeMaeyer, WHO, Chairman
Dr. Samuel G. Kahn, AID, Secretary

INACG Secretariat

Dr. C. O. Chichester
Dr. Paul D. Maugle

OCTOBER 1984

The preparation of this report was made possible by grant #DAN-0227-G-SS-1029-00 to The Nutrition Foundation, which serves as the Secretariat for the International Nutritional Anemia Consultative Group, from the Office of Nutrition of the Agency for International Development.

Table of Contents

1. INTRODUCTION 1

2. IRON NUTRITION 2

3. SPECIMEN COLLECTION AND ANALYSIS 5
   - 3.1. Introduction 5
   - 3.2. Venous or capillary blood 5
   - 3.3. Sample collection 5
   - 3.4. Specimen transport and storage 6
   - 3.5. Principles of quality control 6
   - 3.6. Measurement of hemoglobin concentration 7
   - 3.7. Measurement of serum ferritin concentration 7
   - 3.8. Measurement of serum iron concentration 8
   - 3.9. Measurement of transferrin saturation 8
   - 3.10. Measurement of erythrocyte protoporphyrin 8

4. PILOT SUPPLEMENTATION TRIALS 9
   - 4.1. Introduction 9
   - 4.2. Questions to be answered 9
   - 4.3. Design of pilot supplementation trials
     - 4.3.1. Definition of target population 9
     - 4.3.2. Organizational considerations 9
     - 4.3.3. Number of study groups 10
     - 4.3.4. Method of administering the supplement 10
     - 4.3.5. Initial stratification of population 10
     - 4.3.6. Stratification by hemoglobin concentration 12
     - 4.3.7. Relative number of subjects in treatment and placebo groups 12
     - 4.3.8. Allocation to study groups 12
     - 4.3.9. Sample size 13
     - 4.3.10. Duration of supplementation and frequency of measurements 13
   - 4.4. Analysis of pilot supplementation trial
     - 4.4.1. Comparability of control and treatment groups 13
     - 4.4.2. Methods of analysis 13
     - 4.4.3. Analysis by comparison of mean change in hemoglobin concentration 13
     - 4.4.4. Analysis of covariance 14
     - 4.4.5. Extrapolation to general population 14

5. FIELD SUPPLEMENTATION TRIALS 16
5.1. Introduction 16
5.2. Questions to be answered 16
5.3. Design of field supplementation trials 16
5.3.1. Definition of target population 16
5.3.2. Organizational considerations and choice of population sample 16
5.3.3. The number of treatment groups 17
5.3.4. Sample size 17
5.3.5. Duration of trial and frequency of measurements 17
5.4. Analysis of field supplementation trials 17
5.5. Cost-effectiveness 18
5.5.1. Cost of supplementation 18
5.5.2. Effectiveness 18

6. APPENDICES 19
6.1. Hemoglobin determination 19
6.1.1. Diluent solution 19
6.1.2. Standards 19
6.1.3. Instrument calibration 19
6.1.4. Hemoglobin measurement 21
6.2. Quality control of hemoglobin measurement 21
6.2.1. Accuracy 21
6.2.2. Precision 22
6.3. Quality control of blood collection 24
6.4. Allocation to placebo and treatment groups 26
6.5. Determination of sample size for the comparison of two means 27
6.6. Student's 't' test 28
6.7. One-way analysis of variance 29
6.8. Two-way analysis of variance 32
6.9. Analysis of covariance 34

7. REFERENCES 39

1.
Introduction

This publication is one of a series of manuals published by the International Nutritional Anemia Consultative Group dealing with various aspects of nutritional anemia. It is recommended that this volume should be read in conjunction with the other volumes in the series. As outlined in the first volume, "Guidelines for the Eradication of Iron Deficiency Anemia"[1] ("Guidelines"), in developing a national program for the control of iron deficiency, a combination of strategies such as supplementation, fortification, nutrition education and parasite control will often be necessary. In many countries, however, an important strategy is supplementation, where iron is supplied in medicinal form to vulnerable segments of the population.

In developing programs of supplementation, it is essential that trials of proposed measures be carried out in representative population groups before they are introduced on a regional or national scale. In the last few years, in several countries, trials have been initiated which have highlighted some of the difficulties which arise in the design and analysis of such trials. The present publication is an attempt to provide guidance for future investigators in this area. It should, however, be emphasized that in every case it is important that a competent statistician be associated with the planning, execution and analysis of any trial. This document is not in any way meant to be a substitute for the personal involvement of a statistician.

2.
Iron Nutrition

The healthy, well-nourished individual has a certain amount of reserve or storage iron in the body which may be utilized for metabolic functions in periods of increased iron requirement or decreased iron intake. In periods of continuing negative iron balance the first measurable change in iron nutriture is a progressive decrease in serum ferritin concentration. When iron stores are more depleted, there is no longer adequate iron supply to the marrow. This is reflected in a decrease in the saturation of circulating transferrin, an increase in erythrocyte protoporphyrin and finally a decrease in hemoglobin concentration.

The ultimate aim of intervention strategies is to achieve a state of normal iron nutrition in the target population, or to approach as near as possible to this, within the limitations of available resources. The definition of "normality," however, poses certain problems as outlined in the following paragraphs.

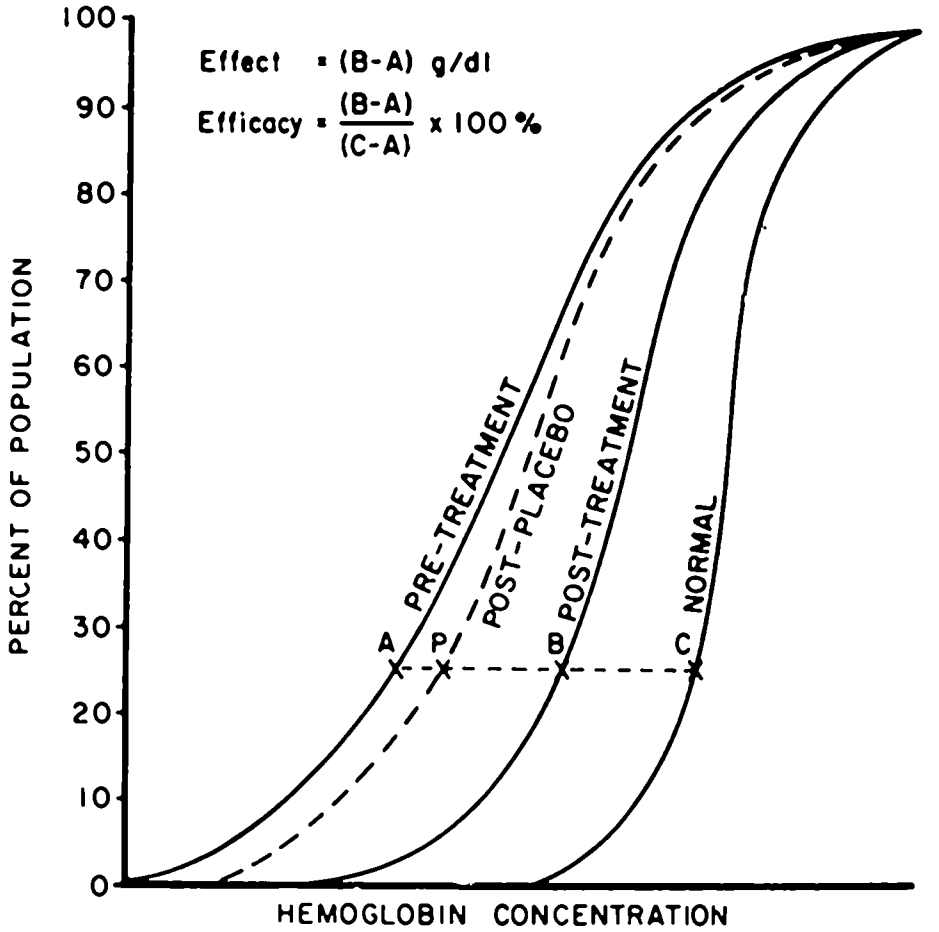
In terms of hemoglobin concentration, the individual has a homeostatic mechanism which, in the absence of disease, and in the presence of an adequate supply of hemopoietic nutrients, sets the hemoglobin concentration at a level which may be considered "normal" for that person.[2] Further supply of hemopoietic nutrients will not increase the hemoglobin concentration. The frequency distribution of normal hemoglobin concentration is influenced by age, sex and altitude of residence. There is also evidence that there may be some racial differences.[3,4] For a population living at a given altitude, the frequency distribution of normal hemoglobin concentration of individuals in the different demographic segments of the community (e.g., adult males, nonpregnant women of childbearing age, preschool children, etc.) approaches closely a Gaussian distribution.

The homeostatic mechanisms do not control body iron stores as precisely as they do the hemoglobin concentration. As iron stores increase, iron absorption is reduced, but not abolished. There are therefore no clearly demarcated limits for "normal" iron stores. The best simple indicator of iron stores in healthy people is the serum ferritin concentration. At birth, geometric mean values are in the region of 200 µg/liter. These fall during the first year of life to about 35 µg/liter. Values in the adolescent female stay at about this level until the end of reproductive life and then rise, while in the adolescent male they rise to a geometric mean of about 100 µg/liter and increase slightly with age.[6] At all ages, values below 12 µg/liter are indicative of iron deficiency and values above 500 µg/liter (in the absence of disease states) are suggestive of iron overload. Values falling within this range may be taken as normal.

With regard to saturation of transferrin, in adults values above 30% suggest an adequate iron nutritional status and values below 15% are suggestive of iron deficiency. In children from the age of six months to adolescence, normal values are lower and those below 10% suggest iron deficiency.[6] Erythrocyte protoporphyrin rises as iron becomes unavailable for heme formation and values above normal (depending on the method employed) are indicative of iron deficient erythropoiesis.

Trials of iron supplementation can be assessed by studying the improvement in hemoglobin and serum ferritin concentrations and other indices of iron nutrition brought about by the intervention as compared with a control situation. This is referred

FIGURE 1



Hypothetical cumulative frequency distribution curves of hemoglobin concentration of a demographic segment of a "normal" population, and a population with a high prevalence of anemia before intervention ("pre-treatment") and after intervention ("post-treatment"). The effect of an intervention can be judged by determining the improvement in hemoglobin concentration at, e.g., the 25th percentile of the population (B - A) g/dl. Similarly, the efficacy of an intervention can be determined by how closely the hemoglobin concentration at the 25th percentile moves toward the normal $\frac{(B-A)}{(C-A)} \times 100\%$. Note, this assumes that there is no change in the placebo group. If subjects given a placebo show an improvement in the frequency distribution of their hemoglobin concentration (dotted line, "post-placebo"), then the effect of the treatment will be (B - P) g/dl, and the efficacy will be $\frac{(B-P)}{(C-P)} \times 100\%$.

to hereafter as the "effect" of the intervention (Figure 1). Trials may also be assessed by seeing how close indices of iron nutrition in the target population come to normality. This is referred to hereafter as the "efficacy" of the intervention (Figure 1). It must be emphasized that "efficacy" can only be judged when the normal is known for the demographic sector of the population under study. In terms of hemoglobin concentration, this knowledge can perhaps be approximated by extrapolation from studies in other geographical areas, but the only definitive way of obtaining the information is by determining the frequency distribution of hemoglobin concentration in a sufficiently large group of apparently healthy individuals representative of that demographic sector, who are either supplemented with all known hematinics until there is no further change in their hemoglobin con-

centration, or who are shown to have adequate stores of all hematinics, especially iron, folate and vitamin B_{12}.

The precise end point of any trial will be dependent on a variety of factors. Ideally, the end point would be to restore the frequency distr bution of hemoglobin concentration of the tarş et population to normality and to ensure that all the population has norma! iron stores. In practice, the limitation of available resources may make this an impossible objective and, at least initially, a more readily achievable end point may have to be set, such as producing a significant reduction in the prevalence of anemia.

3.
Specimen Collection and Analysis

1. Introduction

Since the assessment of any supplementation trial depends on laboratory measurements, it is essential that detailed attention be given to all aspects of the collection, transport, storage and analysis of blood specimens.

The most important single measurement is the hemoglobin concentration. In addition, study of the other indices of iron nutrition -- serum ferritin concentration, serum iron concentration, transferrin saturation and erythrocyte protoporphyrin -- will be likely to provide further information of value, particularly among the less anemic segments of the population. If desired, hematrocrits can also be measured, but since they are rather more difficult to determine accurately in the field situation (due to voltage fluctuations, etc.) and since the hematocrit is not as sensitive an index of anemia, the hematocrit should not replace the determination of hemoglobin concentration.

3.2. Venous or capillary blood

Ideally, all measurements should be made on venous blood samples since, at least in the case of hemoglobin concentration, the results are more reproducible than those on capillary blood. Also, measurements of serum iron concentration and transferrin saturation need more blood than can readily be obtained from a capillary specimen. In many communities, however, there is much greater resistance to venepuncture than there is to the collection of capillary blood. A decision will have to be made initially as to what type of blood specimens are to be used depending on local circumstances. If it is not possible to get venous samples in all subjects, it can be decided either to use capillary blood in all subjects and omit measurements of serum iron and transferrin saturation, or to use capillary blood in some subjects and venous blood in others (e.g., capillary in children and venous in adults). If the latter course is adopted then it is important to ensure that all measurements in a given individual are done on the same type of sample throughout the course of the study. This requires that the blood collector has an appropriate record on every individual. Care must also be taken that there are enough individuals in any subsample selected to allow for valid statistical analysis of the parameters being measured, otherwise nothing is gained by studying such a subsample.

3.3. Sample collection

Since a variety of physiological factors such as exercise, posture, apprehension, etc. may affect blood volume (and hence hemoglobin concentration) it is important, as far as possible, to standardize conditions for blood collection. Specimens should be obtained at the same time of day, subjects should rest sitting down for at least fifteen minutes before blood is collected and blood should be drawn with the subject in a standardized position, either sitting or lying (in the standing position hemoglobin concentration increases due to hemoconcentration).

Venous blood should preferably be obtained either with disposable plastic syringes and needles or with "vacutainers." When measurements of serum iron are to be made, care must be taken to ensure that syringes, needles and containers are iron free. The specimen should be obtained preferably without the use of a torniquet, or with a torniquet applied for less than one minute and released before drawing the blood (when a torniquet is left in situ, local hemoconcentration occurs and this increases

with time); 5 ml of blood should be drawn, 1 ml added to a suitable anticoagulant such as Na_2EDTA (1.5 mg per ml of blood), heparin, or oxalate, for hemoglobin and red cell protoporphyrin determinations and 4 ml allowed to clot for serum ferritin, iron and transferrin determinations.

Capillary blood should be obtained by a clean puncture of fingertip, ear lobe, or, in babies, the heel, preferably with a standardized disposable lancet. Blood should be free-flowing and no squeezing or compression employed as this is liable to dilute the specimen with tissue fluid. Capillary blood for hemoglobin estimation should be obtained by using a standardized "to contain" 20 µl heparinized capillary tube. The tip of the capillary tube is placed in the blood and the tube allowed to fill just beyond the mark by capillary action. The outside of the capillary tube is then wiped with damp gauze, taking care not to discharge the sample from the tube. The amount of blood in the capillary tube is then adjusted exactly to the mark by moving a finger across the distal tip of the tube. The outside of the capillary tube is again cleaned and the contents discharged into the lower part of a measured volume (e.g., 4.0 ml) of the diluent solution. The capillary tube is then rinsed out 8 - 10 times with diluent solution from the upper part of the diluent solution and the whole is securely capped and mixed by inversion. Capillary blood for ferritin estimations is best collected in wide bore capillary tubes (100 × 1.6 mm) which are allowed to clot and subsequently centrifuged and the serum separated.

In order to provide some degree of quality control, it is recommended that, where possible, two independent specimens should be obtained from every fifth or tenth subject sampled by a given worker and the specimens submitted blind to the laboratory. An analysis of the variability of these duplicates, among different workers, will enable some degree of supervision of the quality of work of the field collectors to be maintained (#6.3).

3.4. Specimen transport and storage

Ideally, blood specimens should be collected immediately adjacent to the laboratory. In most field studies, however, this will not be possible and specimens will have to be transported over some distance. This is best accomplished by keeping all specimens at 4°C in an ice box or refrigerated container. Hematocrit measurements, when employed, should be made within two hours of obtaining the blood to obviate artifacts brought about by alterations in red cell volume. Hemoglobin estimations on venous or capillary blood should preferably be made within 24 hours either in a field laboratory, or specimens should be transported to the central laboratory at least once each day while field collections are in progress. Samples for protoporphyrin analysis may be kept at 4°C for up to 8 days, but they should not be frozen, as this affects the extractability of the protoporphyrin. Serum should be separated within 24 hours of blood collection. Sera for iron and transferrin may be stored frozen (-20°C) for up to several months before analysis and samples for serum ferritin may be kept at -20°C for several years, provided they are well sealed in small (e.g., 400 µl) vials with a relatively small air space. Before analysis such stored samples must be thawed and thoroughly mixed.

Where studies are being done in more remote places, it may be possible to do hemoglobin and erythrocyte protoporphyrin estimations on fixed volumes of blood which are pipetted onto filter paper and allowed to dry, transported to the laboratory in a dried state and subsequently eluted.[8] Any such technique, however, must be carefully validated and standardized in the individual laboratory.

3.5. Principles of quality control

In the performances of any trial, detailed attention to assure continued accuracy and precision of all laboratory methods is essential.

The accuracy of a given procedure is defined as the difference between the observed value and the "true" value. Accuracy is monitored by introducing one or more known standards with every batch of specimens run (#6.2.1). If the observed value differs by more than a predetermined amount, the reason must be sought and corrective action taken.

The precision of a laboratory procedure is a measure of its repeatability as influenced by random errors. Precision is defined statistically as the standard deviation (S) in replicate determinations on aliquots of the same sample where

$$S = \sqrt{\frac{\Sigma d^2}{n}}$$

where d is the difference between each value and the mean ($\bar{X}$) of all the values, Σd^2 is the sum of the

squares of all the d's and n is the number of determinations. Alternatively, precision may be expressed as the coefficient of variation (C) where

$$C = \frac{S}{\bar{X}} \cdot 100$$

To monitor precision of a laboratory method a control chart for variation in estimations of aliquots from the same sample should be established prior to beginning the routine testing of specimens (#6.2.2). For every batch of specimens run, a minimum of four aliquots of one specimen should be randomly distributed among routine specimens and an estimate of the variability, S, of these measurements is plotted on the control chart. If any estimate of S falls outside the control limits, the method is out of control and the reason must be determined and the measurements obtained in this batch must be discarded. If material is available, the measurements can be repeated when corrective action has been taken. If any point falls between the warning and control limits, or if seven consecutive points fall on one side of the central line, the reasons must be investigated and corrective action taken if necessary. Where the precision of the estimation varies considerably over the range of expected values, separate control charts will need to be run for different ranges (e.g., serum ferritin and serum iron #3.7 and 3.8.)

3.6. Measurement of hemoglobin concentration

The cyanmethemoglobin method as specified by the International Committee for Standardization in Hematology (ICSH)[9] should be employed. This method depends on the conversion of hemoglobin to cyanmethemoglobin (HiCN) by adding blood to a solution of potassium ferricyanide and potassium cyanide. The HiCN has a maximum absorbance at a wavelength of 540 nm. The concentration of HiCN may therefore be measured in a suitably calibrated spectrophotometer or colorimeter. Details are given in #6.1.

3.7. Measurement of serum ferritin concentration

There is presently no method recommended by the ICSH. All assays depend on immunological procedures. Details of a two-site immunoradiometric assay and a radioimmunoassay are given in Bothwell et al., (1979);[10] Currently, however, the most practical approach for laboratories not performing the assay on a routine basis is to use one of the commercially available kits which has been demonstrated to give satisfactory results. Most kits employ a radioimmune assay and require the use of radioisotope counting equipment. Enzyme linked immunological methods, however, are also becoming available.

There is an ICSH standard available from the Secretary of the ICSH Panel on Standardization of Ferritin*. Commercial kits contain "standards," but in setting up the assay, a laboratory should also obtain the ICSH standard to help reduce interlaboratory variability. Since serum ferritin appears to be quite stable when stored at −20° and since in immunological assays, between batch variations are always considerably greater than within batch variations, it is suggested that all sera collected from a given patient over a period of time should be run in one batch. It may be desired to know initial serum ferritin values and also to know results during the progress of a long trial (particularly in a field fortification trial) to decide whether results are likely to be significant and the trial stopped, or whether it should be continued. It is, therefore, desirable that when blood is collected, three aliquots should be taken for ferritin estimation. One aliquot can then be analyzed to find out the present status and two can be stored for analysis when all the specimens for that subject have been collected. At the final analysis there should be at least two aliquots, each containing enough sera for triplicate estimations, so that if one run is a failure there is a back-up sample. If commercial kits are employed, it is important that for the final analysis at the end of the trial, there are enough kits on hand from the one production lot to complete the study on all the subjects.

Precision is monitored, as in other tests, by blindly assaying multiple aliquots (a minimum of four) from the one sample. Since precision is greater at higher values, three control charts should be established, one each for high, medium and low ranges. The following within-batch coefficients of variation (CV) should be obtained:

Serum Ferritin μg/l	CV
100	< 10%
30	< 15%
10	< 20%

It should be noted that serum ferritin concentrations have a log-normal distribution, therefore, the control charts for S must be calculated on the log-transformed values.

*Currently Dr. Mark Worwood of the Department of Hematology, University of Cardiff, Wales.

3.8. Measurement of serum iron concentration

The method recommended by the ISCH[11] for the estimation of serum or plasma iron depends on liberating the iron from its binding with transferrin and estimating it by complexing with a suitable chromagen such as 4,7-diphenyl-1-10-phenanthroline (bathophenanthroline). The procedure is described in detail in Bothwell et al., 1979[10].

As with hemoglobin measurements, accuracy must be closely monitored by using standards and precision monitored by dividing one sample into aliquots (a minimum of four) and randomly distributing them among the day's samples. Since precision is greater at higher concentrations, three control charts should be established, one each for high, medium and low ranges. The following within-batch and between batch coefficients of variation should be obtained:

Serum Fe µg/l	CV
150	< 5%
75	<10%
25	<30%

3.9. Measurement of transferrin saturation

Since it has so far proved impossible to obtain an adequate reference standard for the measurement of iron binding capacity, there is as yet no method recommended by the ICSH. The most widely used methods add excess iron to the serum sample, unbound iron is then removed by some adsorbent (ion exchange resin, magnesium carbonate or protein coated charcoal) which complexes with unbound iron but not with bound iron, the adsorbent is separated by centrifugation and either the total iron remaining in the supernatant is measured, giving the total iron binding capacity, or, if radioactive iron is used, the amount of radioactivity in the supernatant is a measure of the unsaturated iron binding capacity. The magnesium carbonate method is recommended. This method is described in detail in Bothwell et al., (1979)[10]. Although no satisfactory reference material is available, a locally prepared "standard" can be made by pooling sera from six "normal" iron replete adult males (which should give a value of approximately 300 µg/dl) and the "true" value determined by performing replicate analyses and obtaining the mean. Multiple aliquots of this standard can be kept frozen for many months and one or more thawed and measured in each run of specimens. This procedure will monitor the repeatability of the assay. Precision is monitored, as with other measurements, by dividing one specimen into four or more aliquots and randomly distributing them among a batch of samples. The within-batch and between-batch coefficient of variation should be <10%.

The saturation of transferrin is given as the ratio of the serum iron (SeFe) to the total iron binding capacity (TIBC) expressed as a percentage, i.e.,

$$\% \text{ saturation} = \frac{\text{SeFe}}{\text{TIBC}} \cdot 100$$

3.10. Measurement of erythrocyte protoporphyrin

There are a number of methods available for estimation of erythrocyte protoporphyrin which give satisfactory results. All depend on fluorometric measurements. Perhaps the best way of expressing results is as the erythrocyte protoporphyrin/heme ratio. A suitable method is described by Labbe et al., (1970)[12]. There are presently no suitable standards for establishing the accuracy of the method, but the coefficient of variation is approximately 5% for normal blood and less than this in blood from iron deficient subjects where the protoporphyrin/heme ratio is increased. No systematic between-laboratories standardization of erythrocyte protoporphyrin measurements has been attempted. Each laboratory, therefore, must establish its own normal range.

4.
Pilot Supplementation Trials

4.1 Introduction

In communities where there is a high prevalence of iron deficiency anemia, supplementation programs may be needed to improve the hemoglobin concentration of the more severely affected individuals within a reasonable period of time. This is particularly applicable, for example, to pregnant women when it is desired to correct the anemia and improve the iron nutritional status before the end of pregnancy. Because there is likely to be a high prevalence of moderately severe anemia, the chief parameter of interest will be the hemoglobin concentration.

4.2. Questions to be Answered

As outlined in the "Guidelines,"[1] where there is insufficient knowledge about the response of the target population to therapy, an initial pilot supplementation trial will have to be carried out to answer the following questions:

a) Will supplemental iron improve the hematological status of the target group (i.e., is iron better than placebo)? This will define whether or not there is a limiting deficiency of iron in the target group.
b) Is there a need for any other hematinic in addition to iron (e.g., is iron plus folate better than iron alone)?
c) What is the best formulation, frequency of administration and dose of supplemental iron in terms of therapeutic response, acceptability and side effects?
d) What is the effect and efficacy of the best available supplement?

The pilot supplementation trial will also provide an estimate of the within-treatment variance of the measured parameters which can be used to decide the number of subjects required in future studies.

4.3 Design of pilot supplementation trials

4.3.1. Definition of target population

The target population should be clearly identified. It will usually consist of people living in a delineated geographic area who are at risk of developing iron deficiency anemia and who belong to a specific demographic sector, such as pregnant women or preschool children.

Ideally, pilot trials should be carried out in a randomly selected sample of the target population. In practice this sampling may be difficult, but much information can be obtained by studying a non-randomly selected sample such as women in the first or second trimester of pregnancy attending an antenatal clinic. It must be emphasized that results obtained from the study of such a selected sample can only be extrapolated to the target population to the extent to which the clinic population is representative of the whole.

4.3.2. Organizational considerations

The organizational details will differ depending on the target population and the local situation.

In all cases adequate laboratory facilities are essential. It is desirable that the analysis be entrusted to a special qualified technician, or technicians, given suitable training to ensure that the highest standards of laboratory practice are maintained at all times. Where the trial is conducted away from the

laboratory, appropriate arrangements must be made for transport of specimens from the site of collection to the laboratory.

The proposed trial should be cleared by the ethical committee of the Investigator's Institute. Where no ethical committee exists, it might be wise to form an ad hoc ethical committee, consisting of prominent lay and medical personnel not concerned with the study, to supervise the ethical considerations involved in the establishment and conduct of the trial.

The trial must be fully explained to the target population and their informed consent obtained before they are included in it.

Suitable arrangements must be made for data recording and analysis. If numbers are relatively small, this may be done by hand. If the numbers involved are large, however, mechanical or electronic data processing will be necessary.

4.3.3. Number of study groups

The number of study groups will depend on the objectives of the pilot supplementation trial. The simplest situation is a pilot trial to determine whether iron is better than a placebo, carried out in a target population in which there are no independent variables which may affect the outcome. Such a trial requires two study groups — one treatment and one placebo. Often it is possible to design the pilot trial to answer several of the questions in #4.2 at the same time; this will require one group for each supplementation strategy and one placebo group. Care should be taken, however, not to make the trial too complicated. Experience suggests that there should not be more than six treatment groups plus a placebo group, and preferably less than this.

4.3.4. Method of administering the supplement

To eliminate differences due to variability of supplement consumption, pilot trials should usually be carried out with close supervision to ensure that the supplement is actually swallowed. Careful records should also be kept of the supplement consumption of each individual. At times it may also be desirable to include a group in which administration of the supplement is unsupervised to test a possible "real-life" delivery system.

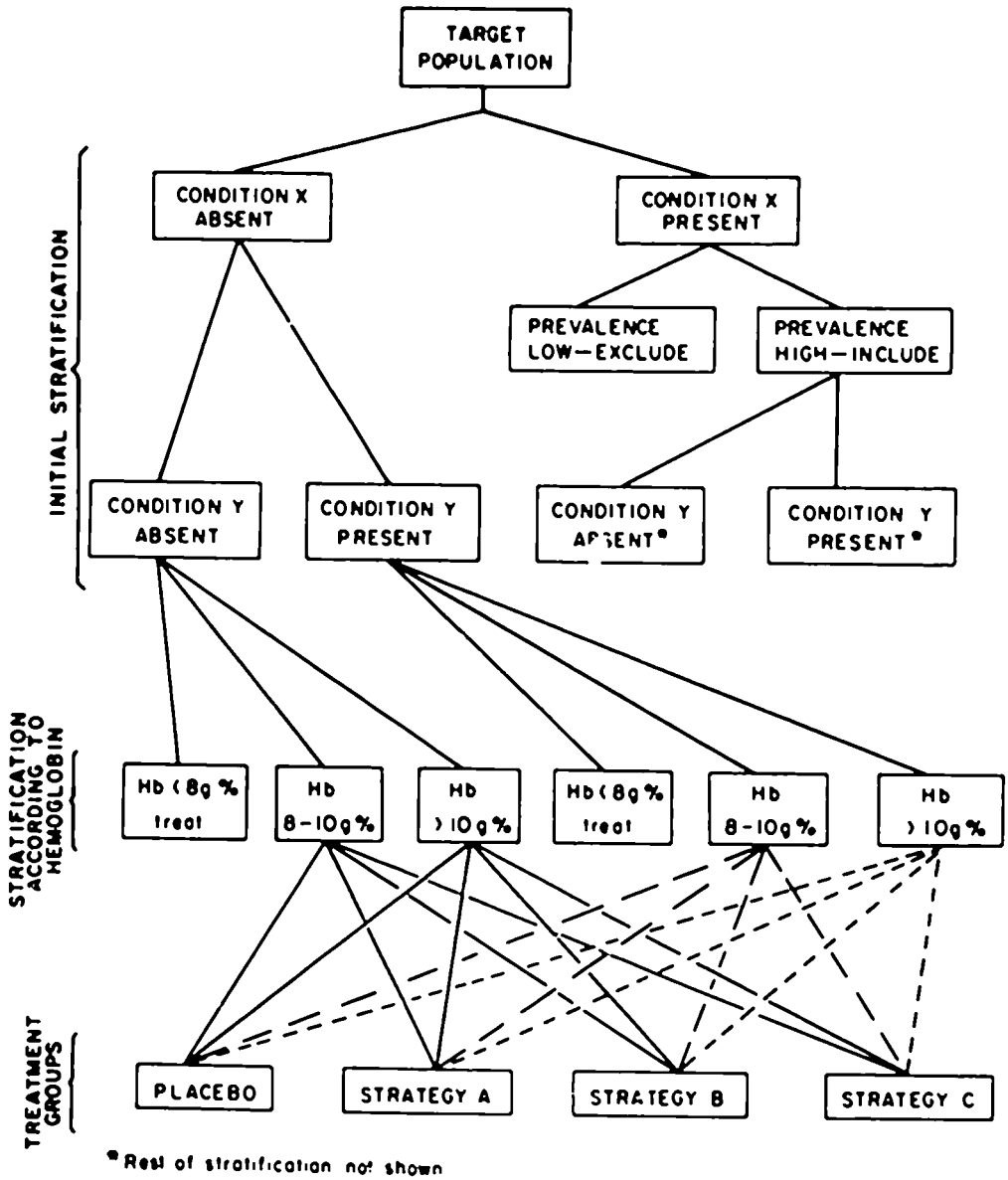
The subjects, the supplement administrators and the laboratory workers responsible for sample analysis should not know which therapy is being administered or which group the subject belongs to. The placebo and the various supplements should be identical in appearance (size, shape, color, number, etc). The supplement for each subject should be prepackaged (e.g., for a week at a time) by someone who has the key, and each package clearly labelled with the patient's identification, but not carrying any indication as to the treatment group. Obviously in some situations, for example, when it is desired to examine the difference between once daily or twice daily supplement administration, it cannot be a completely blind trial. In such a situation, half the placebo group should receive the "supplement" once a day and the other half should receive it twice a day, thus preserving an element of blinding.

4.3.5. Initial stratification

Steps should be taken, as far as possible, to ensure the homogeneity of the various groups. Factors prevalent in the target population which may affect the hemoglobin concentration, or influence the response to therapy, should be defined. Individuals with chronic diseases and infections such as tuberculosis, leprosy and malaria should be excluded because of the difficulty in interpreting results. Those with thalassemia should be excluded because of the risk of iron overload. Where there are many individuals with conditions such as hemoglobinopathy and hookworm infestation, they must be identified and initial stratification done according to the presence or absence of these conditions. An example of such initial stratification where hemoglobinopathy and hookworm infestation are prevalent is shown in Figure 2. Since the presence of certain hemoglobinopathies* (condition X, Figure 2) may affect hemoglobin concentration irrespective of iron nutrition, the division into X absent and X present is done first. If there is a low prevalence, subjects with this condition can be excluded, but where there is a high prevalence, it may also be important to study the effect of supplementation in these individuals. Two streams must therefore be established — one consisting of individuals without X and the other of individuals with X, and these streams must be maintained throughout the study.

*Studies in Burma (Aung Thang Batu, Personal Communication) and Thailand (Aree Valyasevi, Personal Communication) have shown that the response to iron therapy of subjects with hemoglobin AE is closely similar to that of those with hemoglobin AA.

FIGURE 2



Stratification of target population with two factors other than hemoglobin concentration that may affect the outcome. The full stratification for X present—prevalence high is not shown, but is similar to the left hand side X absent.

In many countries where trials may be carried out, hookworm infestation may be prevalent. For a given species of hookworm the daily additional iron loss from the body is directly proportional to the number of parasites. It is therefore desirable to ensure that both the number of affected individuals and the frequency distribution of the parasite load among affected individuals are as uniform as possible in the different treatment groups. For the purposes of stratification, the best dividing line will depend on the prevalence and seventy of hookworm infestation. Where loads are light, the best division may be between those who have hookworm and those who don't (condition Y, Figure 2).

In other instances, where there is a high prevalence of heavy infestation, a division into two streams of those with less than 2000 eggs per gram of feces and those with 2000 or more[13] may give a more even distribution of individuals with heavy hookworm loads between the different treatment groups.

4.3.6. Stratification by hemoglobin concentration

In order to ensure reasonable homogeneity of the groups, the streams defined by the initial stratification should be further subdivided according to the hemoglobin concentration. Individuals with a hemoglobin concentration below a certain level will have to be excluded for ethical reasons and treated. Those with hemoglobin concentrations above this level should be suitably stratified. In practice, where there is a high prevalence of anemia and individuals with a hemoglobin less than 8 g/dl are excluded, division into two strata consisting of individuals with hemoglobin from 8 to 10 g/dl and those with hemoglobin over 10 g/dl is often satisfactory. Each person, in each of the above hemoglobin streams, is then numbered consecutively for random allocation to a study group.

4.3.7. Relative number of subjects in treatment and placebo groups

When more than one treatment group is to be compared with a control (placebo) group, theoretical considerations require that the number of subjects in the control group is greater than the number in each treatment group. This is given by the formula:

$$c = t\sqrt{m}$$

where c is the number of subjects in the control group, t is the number of subjects in *each* treatment group and m is the number of treatment groups (i.e., excluding the control or placebo group).

4.3.8. Allocation to study groups

Allocation to study groups is most simply done by the use of random permutation tables either taken from published tables or constructed for the purpose (#6.4.).

As an example, consider individuals in the 8 to 10 g/dl hemoglobin stream (condition X absent, condition Y absent) who are to be allocated in equal numbers (t) to each of three treatment strategies (A), (B) and (C) (Figure 2). The above formula then requires that there should be $t\sqrt{3}$ in the placebo group. This means that for every 3 subjects in each treatment group there should be $3 \times \sqrt{3} = 3 \times 1.73 \approx 5$ in the placebo group, i.e. for every nine individuals (3 × 3) in the treatment groups there will be 5 in the placebo group. Random permutation tables of blocks of fourteen numbers (9 + 5) are therefore needed. Two such permutations of fourteen numbers are shown below:

1,	6,	8,	5,	10,	12,	11
9,	2,	3,	14,	7,	4,	13
11,	3,	8,	10,	4,	9,	6
14,	5,	2,	7,	13,	1,	12

The symbols for the placebo (P) and three treatment groups (A), (B), and (C) are written in the following order

PABC, PABC, PABC, PP

under the first block of numbers thus:

1,	6,	8,	5,	10,	12,	11
P	A	B	C	P	A	B
9,	2,	3,	14,	7,	4,	13
C	P	A	B	C	P	P

The random numbers then represent the serial number of the subject within that hemoglobin stream, so that subject no. 1 is allotted to placebo (P), subject no. 2 to placebo (P), subject no. 3 to strategy A, etc., and subject 14 to strategy B. When the first block of random permutation numbers has been used, the next block is used in an identical fashion, the "1" in the block now standing for serial number 15, 2 for 16, etc. and 14 standing for serial number 28. This process is continued with successive blocks until there are enough serial numbers to cover all subjects required in that hemoglobin stream. A similar process is employed with the allocation of individuals in the other hemoglobin stream.

The cyclic order of allocation for other numbers of treatments, the size of the random permutation blocks required and the method for constructing random permutation blocks are given in #6.4.

In order to avoid any possible bias in allocation of subjects, it is important that random allocation be prepared in advance and placed in sealed envelopes by someone who will not be involved in subject allocation. Each envelope should bear on the outside the identification of the stream (e.g., 8 to 10 g/dl hemoglobin stream, condition X absent, condition Y absent) and the serial number of the subject and sealed inside should be the same identification data plus the study group to which that subject is

allocated. The envelope is then opened only after the decision has been made to admit the subject to the study.

4.3.9. Sample size

The sample size required for analysis depends on the variability of the parameter of interest and the minimum change it is desired to detect at a given level of probability (power) and a given confidence coefficient. In supplementation trials, the major parameter of interest is the hemoglobin concentration. The method of calculation of sample size required for the comparison of two means is given in #6.5 together with a table indicating sample size based on an estimated standard deviation (σ) of hemoglobin concentration in the population of 1.0 and 1.2 g/dl, a confidence coefficient of 95% and powers of 90% and 80%.

The initial sample size must be greater than the number required for analysis to allow for dropouts due to lack of cooperation, migration, illness, etc. When recruitment to the trial is done longitudinally (e.g., women attending an antenatal clinic), dropouts can be compensated for by continuing recruitment for an appropriate period of time until adequate numbers of subjects have completed the trial. When recruitment of the sample is done at one point in time (e.g., school children), however, an appropriate increase in numbers must be made at the commencement of the study.

At times, sample size may be high in relationship to the resources (money, time, personnel) available for the trial. If this is so, it may be possible to confine the study to the most anemic segment of the population in whom the greatest change in hemoglobin concentration can be anticipated; but if this is done, the results will not necessarily be applicable to all the population of interest.

4.3.10. Duration of supplementation and frequency of measurements

The duration of the trial and the timing of measurements will depend to some extent on the target population. In many areas of the world, pregnant women may only be available for supplementation for a period of 10 to 14 weeks. In this case, measurements may be made at the beginning and end of supplementation, the latter measurement clearly must be made before delivery. In other instances (e.g., preschool children), supplementation may be carried out for a number of months and measurements made at 3 to 6 month intervals. The intermediate measurements will provide information on the secular trends of hemoglobin concentration. In all cases, the duration of study and frequency of measurements must be the same in the placebo and each of the treatment groups. In the case of supplementation carried out over a prolonged period of time, in order to avoid possible seasonal influences, it is important that measurements in the placebo and each of the treatment groups be made at the same time of the year.

4.4. Analysis of pilot supplementation trial

4.4.1. Comparability of control and treatment groups

The first step is to check that the random allocation procedure, after excluding dropouts, has resulted in comparable test and control groups in terms of initial hemoglobin concentration and relevant associated variables such as hookworm infestation, etc. In the case of hemoglobin concentration, if there are only two groups, this may be done by the Student's 't' test (#6.6) or where there are more than two groups, by a one way analysis of variance comparing the mean initial hemoglobin concentration of the groups (#6.7). In the case of hookworm, the distribution of egg counts per gram of feces is likely to be log-normal. If so, a logarithmic transformation of the egg counts should be done and the comparison of the hookworm loads done by a Student's 't' test if there are two groups, or by a one way analysis of variance, if there are more than two groups.

4.4.2. Methods of analysis

The more anemic an individual, the greater is the response which may be produced by supplementation. In analysis of responses, it is therefore important to allow for possible small differences in the initial hemoglobin concentrations in the different study groups. This can be done in two ways — either by analysis of the change in hemoglobin concentration in individual subjects or by analysis of covariance.

4.4.3. Analysis by comparison of mean change in hemoglobin concentration

a) When placebo and treatment groups are initially comparable.

If the placebo and treatment groups are shown to be initially comparable (#4.4.1), analysis may be made by comparing the mean change in hemoglobin concentration in the different groups. For each individual, the change in hemoglobin concentra-

tion "d" (d = final - initial) is computed. Within each group, the overall "mean d" ($\bar{d}$) is calculated.

When there are only two groups, one placebo ("P") and one treatment ("T"), the statistical significance of the overall effect of treatment may be assessed by comparing "$\bar{d}_P$" and "$\bar{d}_T$" by a Student's 't' test (#6.5). If, however, there are more than two groups, the mean d's in the different groups may be compared using a one-way analysis of variance (#6.7).

In addition to the comparison of the overall effect of treatment, it is of interest to compare the mean change in hemoglobin concentration in the different treatment groups within the two initial hemoglobin streams (#4.3.6), 8 to 10 g/dl and <10 g/dl, i.e., $\bar{d}_{(8-10)}$ and $\bar{d}_{(<10)}$. When there are only two treatment groups, the mean change within each hemoglobin stream may be compared using the Student's 't' test, i.e., $\bar{d}_{(8-10)P}$ is compared with $\bar{d}_{(8-10)T}$ and $\bar{d}_{(<10)P}$ with $\bar{d}_{(<10)T}$. If, however, there are more than two groups, the comparison of mean d's has to be made using a two-way analysis of variance (#6.8).

b) When placebo and treatment groups are not comparable initially

If the placebo and treatment groups are not comparable with reference to the initial hemoglobin concentration, it is not valid to assess the statistical significance of the overall effect of treatments by comparing the mean d's. Instead, it is advisable to stratify the individuals in the placebo and treatment groups into 8 to 10 g/dl and <10 g/dl class intervals, or, if need be, into one gram class intervals according to their initial hemoglobin concentration and calculate the d for each of the strata in the placebo and treatment groups. These $\bar{d}$'s within each class interval may then be compared using the Student's 't' test if there are only two groups, or a one-way analysis of variance (#6.7) if there are more than two groups.

In areas where heavy hookworm infestation is prevalent, it may be desirable to substratify within each hemoglobin class interval according to hookworm load (e.g., egg counts of 0 to 2000 and >2000 g of feces) and make the comparison of mean changes using a three-way analysis of variance.* Care must be taken, however, that numbers in the substrata do not become too small for meaningful comparisons.

*Competent statistical advice may be sought to carry out this analysis.

4.4.4. Analysis of covariance

This technique makes comparisons between the final hemoglobin concentrations in the different groups by mathematically making allowances for any differences in the initial hemoglobin concentration in the groups (#6.9). This is done by fitting a regression line of the final hemoglobin concentration on the initial hemoglobin concentration. This analysis assumes a linear relationship between the initial hemoglobin concentration and the response to supplementation. If this is not true (i.e., if there is a curvilinear relationship), competent statistical advice should be sought regarding the method of analysis. The results may lead to the following general situations:

a) Both the slopes and intercepts of the regression lines for the different groups do not differ statistically. This indicates that there is no significant difference in response between the groups. (Figure 3a)

b) The slopes do not differ, but the intercepts show significant differences. In this case, the differences in intercept reflect treatment differences which are constant over the range of initial hemoglobin concentrations in the sample. (Figure 3b)

c) Both the slopes and intercepts differ statistically among the groups. This reflects significant differences in treatment effect which are not uniform with reference to the initial hemoglobin concentration (Figure 3c)

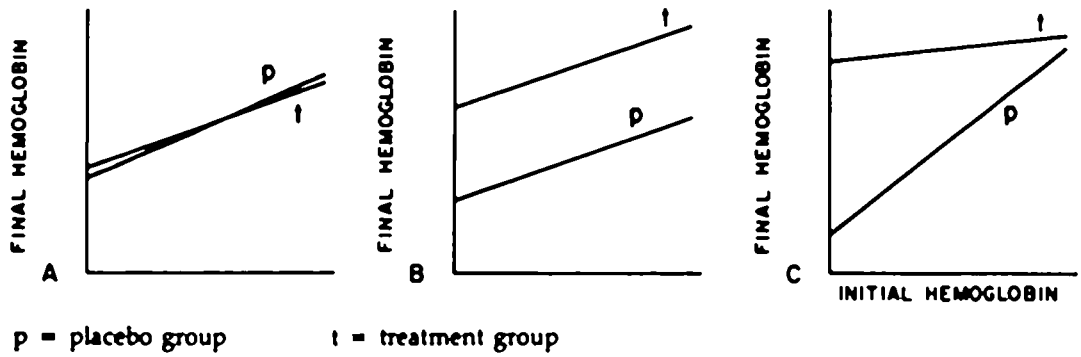
It must be pointed out that in the event of a treatment being completely successful (i.e., all individuals reach their "normal" hemoglobin concentration), the correlation between the final and initial hemoglobin concentration will tend toward zero. In such a situation the superiority of this treatment is obvious.

Where hookworm infestation is prevalent, a covariance analysis using multiple linear regression may be done. Alternatively, separate regression analyses may be carried out for those with hookworm infestation and those without, or those with heavy infestation and those with light infestation (#4.3.4).

4.4.5. Extrapolation to general population

If the subjects studied are not adequately representative of the target population, particularly with respect to variables such as socioeconomic status,

FIGURE 3



Possible results from analysis of covariance of final upon initial hemoglobin concentration (see text).

diet and parasitic infestation, it may be desirable to carry out another pilot trial in a population sample more representative of the target group to ensure that the results obtained initially are indeed applicable to that group as a whole, e.g., if original trials were carried out in an urban hospital-attending population, it might be advisable to repeat the trial in a representative rural setting.

5.
Field Supplementation Trials

5.1. Introduction

In communities where there is a high prevalence of anemia, once the existence of a deficiency has been confirmed and the best form of supplemental strategy determined by a pilot trial as in #4, a field supplementation trial may need to be carried out under realistic conditions of public health practice, to choose the best available delivery strategy. In some situations, it may be possible to avoid the necessity for this step by an appropriate modification in the design of the pilot trial. For example, if the pilot trial includes a group given supplements in an unsupervised fashion, in which the results are as good as in a supervised group, then, clearly, the unsupervised method of distribution is satisfactory.

5.2. Questions to be answered

Field supplementation trials should be designed to answer the following questions:

a) What is the best delivery system for the supplement within the limits of available resources? This may include not only methods of supplement distribution but also programs of education to improve motivation, etc.

b) What is the effect and efficacy of using this delivery system?

c) What is the cost effectiveness of the intervention?

5.3. Design of field supplementation trials

5.3.1. Definition of target population

The target population is the same as that for pilot supplementation trials. In this case, however, it is essential that the trial be carried out in as representative a sample of the target population as possible.

5.3.2. Organizational considerations and choice of population sample

A field trial is a much bigger undertaking than a pilot trial. The organizational considerations for a pilot trial (#4.3.2) apply, but many more resources are needed because larger numbers and separated geographical communities will usually be involved.

At all times, close liaison must be maintained with the community leaders, the health authorities, private medical practitioners and other health workers whose cooperation in the study must be assured.

Ideally, the different treatment groups should be made up of individuals randomly selected from the target population in a given geographic area. However, if it is not possible to limit the effects of a given strategy to only a portion of the target population in a given area, different strategies may have to be tested in separated geographical units. If this is so, then, in order to permit analysis of between-unit variations, it is necessary to use *each* strategy in subjects in at least three different units.

In the allocation of the different strategies to different geographical units, it is important to maintain as even a balance as possible with respect to the frequency distribution of hemoglobin concentration and other parameters which may affect the outcome. In order to do this, it will probably be necessary to make an initial hematological survey of a randomly selected subsample of the demographic sector of interest in the various geographical

units to determine the prevalence of anemia and of confounding variables.

5.3.3. The number of treatment groups

The number of treatment groups will depend on the number of strategies being investigated. A placebo group must be included to ensure that the observed effects are due to the administered supplement. Ideally, a group with supervised supplement administration should also be included. In practice, however, where the pilot trial was carried out in a reasonably representative population sample, it may be possible to assume that the result of supervised administration is likely to be closely similar in the field situation; in this case, a supervised group may be omitted.

5.3.4. Sample size

The principles of the calculation of sample size are the same as those outlined for a pilot supplementation trial (#4.3.9). In a field trial, the major contrast of interest is the differences in response between groups receiving different strategies, rather than the difference in response between placebo and treatment groups. These differences are likely to be of smaller magnitude than those between the placebo and the other groups, therefore, it is desirable to choose a sample size which will permit the detection of these smaller differences. The precise d which it is desired to detect may vary according to local circumstances and available resources, but a difference in response of less than 0.5 g/dl in hemoglobin concentration may be of doubtful overall importance.

Where there are factors in the target population which may influence the response to therapy, numbers will have to be increased appropriately to allow for post-stratification procedures. For example, if hemoglobinopathy* is prevalent and it is desired to study the effect of supplementation in these subjects, as well as in those without this condition, then adequate numbers of both types of subjects must be included in the original sample so that, at the time of analysis, the required minimum numbers will be available to provide satisfactory analysis of those individuals with and without hemoglobinopathy. Similar considerations would apply to any other condition which may materially affect the outcome.

When each strategy and the placebo must be studied in different geographical units (#5.3.2), the number given in Table 8 (#6.5) is the minimum number of subjects, in each demographic sector of interest in the population, who must be present *in each geographical unit*, without taking into account the correlation that is likely to occur between individuals within a cluster and the cluster size. Since estimates of the intraclass correlation and cluster size are likely to vary in different study areas, it is not possible to make precise provision for this in determining sample size. As a rule of thumb, it may be desirable to double the sample size shown in the table to allow for this. As in any trial, appropriate allowance must also be made for dropouts.

When the number of individuals in each demographic sector in each unit is close to the number required, all of the individuals should be included in the investigation. If, however, the number of individuals in a given demographic sector in a unit is considerably larger than the number required, then an appropriate random sample should be chosen for the investigation and the same individuals followed throughout the study (i.e., a panel sample).

5.3.5. Duration of trial and frequency of measurements

These will be the same as for pilot supplementation trials.

5.4. Analysis of field supplementation trials

Unlike pilot trials, pre-stratified field trials designed to improve initial homogeneity of the groups may not be feasible. In such circumstances, it is unlikely that the various groups will be closely similar with respect to frequency distribution of initial hemoglobin concentration. Therefore, analysis of results will have to be done either by post-stratification into 1 gram class intervals according to initial hemoglobin concentration and comparing the mean change in hemoglobin concentration (d) in each class interval by appropriate analysis of variance (#4.4.3), or by analysis of covariance (#4.4.4).

When it has not been possible to perform the trial in one geographical unit, the results of each strategy in each unit must be analyzed separately. If the response to a given strategy is similar in two or more units, the results may be pooled, but, if there is a significant difference in response to the same strategy in different units, then they will have to be analyzed separately and the reason for the difference investigated.

---

*See footnote - page

5.5. Cost-effectiveness

5.5.1. Cost of supplementation

The "cost" of supplementation includes the price of the medication, the cost of its transportation and distribution to the points of administration and the cost of publicity needed to ensure compliance. If administered by existing public health personnel, e.g., in antenatal clinics, the only additional expense will be the staff required for monitoring compliance and effectiveness. If, however, special staff are recruited for tablet distribution, expenses will be much greater.

5.5.2. Effectiveness

The effect of the supplementation on individuals, at each initial hemoglobin concentration, is given by the difference between the post-treatment hemoglobin values in the treatment and placebo groups as estimated by regression analysis. The standard error of the treatment effect at each initial hemoglobin concentration can be calculated from a knowledge of the standard errors of the expected post-treatment hemoglobin value in the treatment and control groups. The latter are obtained from an appropriate analysis of variance.

A visual representation of the effect and efficacy of a given treatment, in any demographic sector of the population, can be obtained by plotting the cumulative frequency distribution of the final hemoglobin concentration of the subjects who received the supplement and comparing it with the frequency distribution of hemoglobin concentration of the subjects who received placebo (effect) and with the normal frequency distribution (efficacy).

If a numerical index of efficacy is desired, the shift in the hemoglobin concentration at a given percentile of the population (e.g., 25th%) as shown in Figure 1 may be employed. Alternatively, the proportion of individuals with hemoglobin concentrations below a limit arbitrarily defining the presence of anemia may be calculated; suggested values are shown in Table 1.

TABLE 1.

HEMOGLOBIN CONCENTRATIONS BELOW WHICH ANEMIA IS LIKELY TO BE PRESENT IN INDIVIDUALS RESIDING AT SEA LEVEL[6]

Age (years)		*Hemoglobin (g/dl)*
0.5 - 4		11.0
4 - 8		11.5
8 - 11		12.0
11 - 14	female	12.0
	male	12.5
14 - 18	female	12.0
	male	13.8
18 +	female	12.0
	male	14.0

6.
Appendices

6.1. Hemoglobin determination

6.1.1. Diluent solution

The diluent solution is made by weighing out 200 mg of potassium ferricyanide, 50 mg of potassium cyanide (both analytical grade) and adding distilled water to make 1 liter. When blood is added to this solution, the reaction needs about ten minutes for completion. This time can be shortened to about 3 minutes by the addition of 140 mg of potassium dihydrogen phosphate and a detergent such as "Nonidet P40" (Shell Chemical Co.) or Triton X-100 (Rohm and Haas) in the amount of 1 ml to each liter. Commercial reagent pellets are also available for making up suitable diluent solutions. The diluent solution should be clear and pale yellow in color. When measured in a spectrophotometer (or colorimeter) at a wavelength of 540 nm, it should have a zero absorbance compared with a water blank. If stored in a dark glass or polythene bottle, this solution is stable for several months. It should not be frozen. If the reagent becomes turbid, has an absorbance other than zero at 540 nm, or a pH outside the 7.0 to 7.4 range, it should be discarded.

Potassium cyanide is very poisonous. It should be stored in a locked cupboard and handled only with the greatest care. The dilute solution is less dangerous since several hundred milliliters would have to be drunk to produce any ill effects.

6.1.2. Standards

Standard reference solutions of HiCN, which conform to the specification of the ICSH, are commercially available. These contain between 55 and 85 mg of Hb per dl and are packaged in ampoules. The solution in these ampoules is stable for years at 4°C, but, once opened, care should be taken to avoid bacterial contamination which would produce turbidity. It is therefore usually recommended that the contents be discarded at the end of the day. The solution should always be allowed to come to room temperature before being used.

6.1.3. Instrument calibration

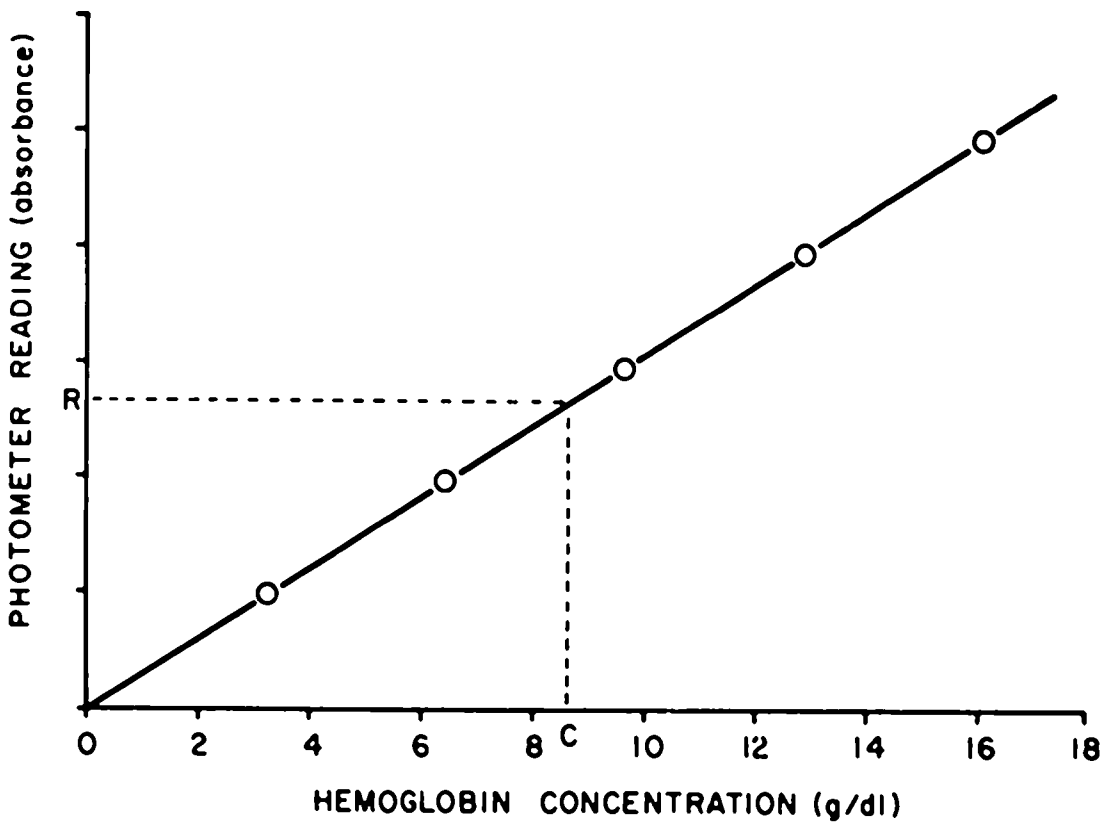
Measurement may be made either with a spectrophotometer using 540 nm wavelength, or a colorimeter using a suitable yellow-green filter covering the range of 530 to 555 nm. Since the electricity supply in many countries may show considerable voltage fluctuations, a double beam instrument which is not affected by moderate voltage fluctuations is advantageous. Alternatively, or in addition, a suitable voltage regulator may be used. Another solution is to use a battery powered instrument.

The instrument should be calibrated by making dilutions in triplicate of a given hemoglobin standard. The following is a suggested protocol, but the actual volumes required will depend on the minimum volume needed for the cuvette of the spectrophotometer or colorimeter being employed.

TABLE 2.
PROTOCOL FOR CALIBRATION OF SPECTROPHOTOMETER OR COLORIMETER

Tube no.	1	2	3	4	5	6
HiCN std. ml	5	4	3	2	1	0
Diluent ml	0	1	2	3	4	5

FIGURE 4.



Calibration curve (absorbance versus hemoglobin concentration) obtained using an HiCN standard diluted as described in the text. The hemoglobin concentration of any unknown specimen is obtained by observing the point at which the photometer reading R intercepts the calibration line and extrapolating this to the horizontal axis C

The volumes should be measured out using the same calibrated 1 ml pipette and each should be done in triplicate. If the standard HiCN solution contains 80 mg/dl of hemoglobin and the measurements are to be made using a 1 in 201 dilution of blood, then the readings obtained with the undiluted standard represent a hemoglobin concentration of 80 × 201 mg/dl or 16.08 g/dl, and the other tubes $\frac{16.08 \times 4}{5}$, $\frac{16.08 \times 3}{5}$, $\frac{16.08 \times 2}{5}$, $\frac{16.08 \times 1}{5}$ and 0, respectively.

The absorbance of each of these should be measured against a reagent blank, the readings taken at each dilution should be averaged and the averages plotted on arithmetic linear graph paper. This should produce a straight line which passes through the origin (Figure 4). Alternatively, and more precisely, all of the individual values can be used and a linear regression analysis performed.

Calibration of the instrument should be carried out at regular intervals (eg. every three months) and whenever it is repaired. In trials where a group of individuals are sampled once a year, such calibration is essential annually before the start of specimen analysis.

The performance of the instrument, once calibrated, should be checked each time before use and at intervals throughout the day. This can most simply be done by having a neutral grey glass standard which is used in place of a cuvette and which produces a known reading on the instrument. Alternatively, a volume of HiCN standard solution, sealed under sterile conditions in a matching cuvette, can be used in the same way. This solution should be kept in the dark and refrigerated when not in use, but allowed to come to ambient temperature before being used in the instrument.

6.1.4. Hemoglobin measurement

Blood samples which have been refrigerated should be allowed to warm up to room temperature. When venous blood is employed, it is essential to ensure thorough mixing of the blood specimen before the aliquot is taken for hemoglobin estimation. This is best accomplished by placing it on a commercially available mixing apparatus for 15 minutes.

20 µl of blood is measured, preferably by using a standardized automatic pipette, and added to 4 ml of diluent (providing a 1 in 201 dilution). This is mixed and time (5 minutes) is allowed for completion of the reaction. When capillary blood is used, dilution will have already been done at the time of specimen collection. The absorbance of the samples is then measured against the diluent as a blank in a matching cuvette. After reading each sample, it is important to ensure that the blank continues to show zero absorbance. The hemoglobin concentration of the sample is then determined by reference to the calibration curve (or a chart derived therefrom) showing photometer readings versus hemoglobin concentration (Figure 4).

6.2. Quality control of hemoglobin measurement

6.2.1. Accuracy

Ideally, all steps in the measurement of hemoglobin should be monitored for accuracy by using a standard indistinguishable from the specimens being measured. This ideal is, unfortunately, not attainable, as whole blood is not stable. Commercially available HiCN standards are already diluted. One or more such standard(s) should be measured each time any unknown specimen is measured, however, this checks only the accuracy of the instrument reading. To check other steps, including pipetting and the suitability of the diluent solution, "local" standards should be prepared as outlined below and at least one run with each batch of estimations.

Local standards may be prepared by withdrawing a volume of blood (e.g., 500 ml) by venesection into acid-citrate-dextrose. Under sterile conditions the plasma is separated and replaced by sterile distilled water. This mixture is then frozen and thawed to ensure complete hemolysis and the red cell ghosts are removed by filtration. Sterile distilled water is then added to produce a hemoglobin concentration of about 11 g/dl. Aliquots of approximately 0.5 ml volume are then put into suitable containers which are sealed, labelled and frozen until used. Provided sterility is maintained, these are stable for at least a year.

Using this standard, a quality control chart to monitor accuracy is set up as follows. Before the commencement of specimen analysis a single aliquot of the standard is analysed once a day for 15 to 20 days, and the mean ($\bar{X}$) and the standard deviation (S) of the estimated values calculated using the formula:

$$S = \sqrt{\frac{(\Sigma X_i - \bar{X})^2}{n-1}}$$

where X_i is the estimated value on the ith day and n is the number of days. An example follows:

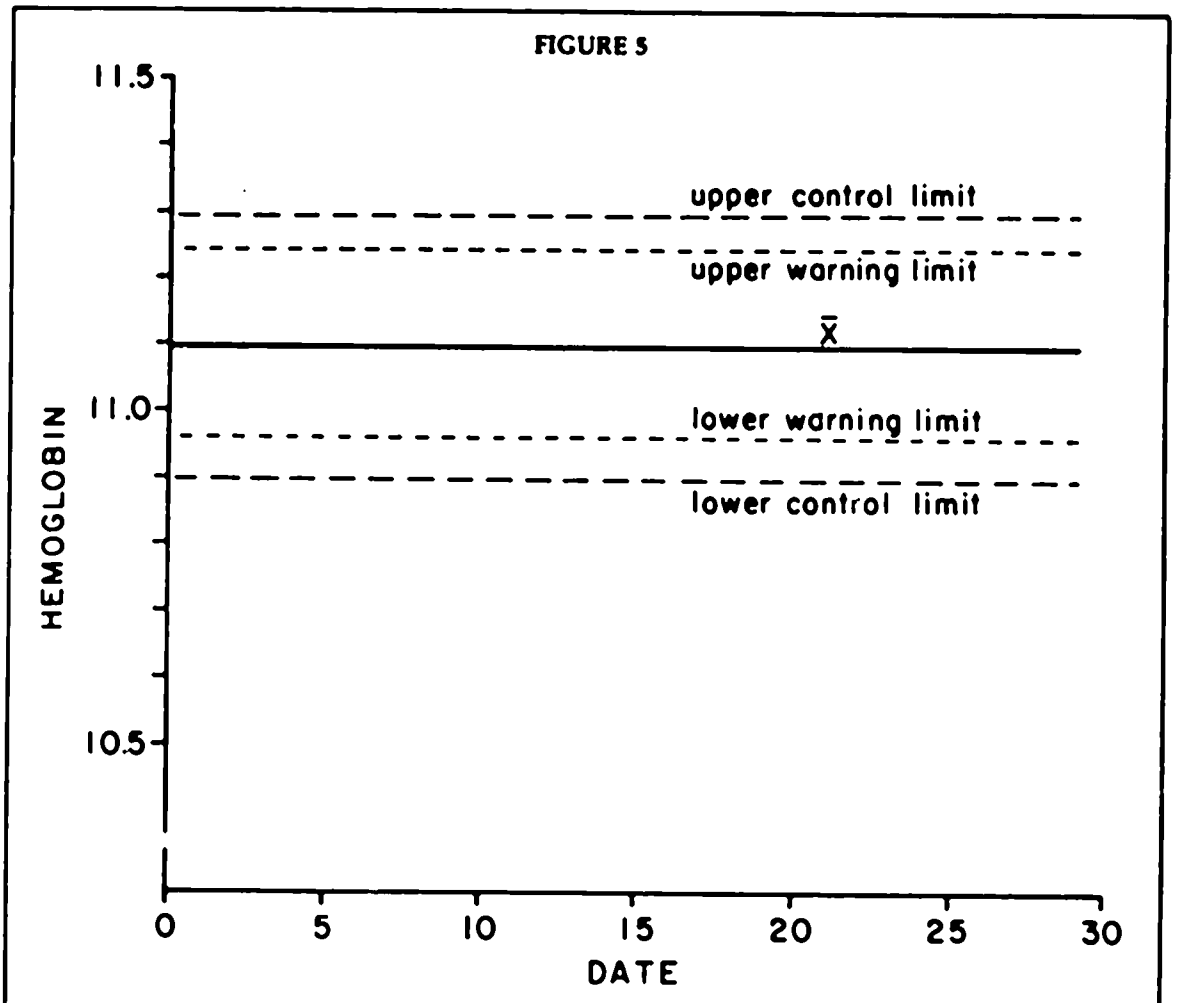
TABLE 3.
DATA FROM REPEATED MEASUREMENT OF HEMOGLOBIN CONCENTRATION OF A "STANDARD"

Day	*X,Hb (g/dl)*	*Day*	*X,Hb (g/dl)*
1	11.1	11	11.1
2	11.1	12	11.2
3	11.1	13	11.1
4	11.2	14	11.0
5	11.1	15	11.1
6	11.0	16	11.1
7	11.1	17	11.2
8	11.1	18	11.0
9	11.0	19	11.1
10	11.2	20	11.2

then $\bar{X} = 11.095$

$$S = \sqrt{\frac{(\Sigma X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{0.0895}{19}} = 0.0686$$

A quality control chart is then drawn with $\bar{X}$ as the central line ± 2S (± 0.14) as the warning limits and ± 3S (± 0.20) as the control limits (Figure 5). Each day, with each batch of specimens, one aliquot of this standard should be introduced among specimens coming from the field. Plotting the estimated value of the standard on the chart on a daily basis will help monitor the accuracy of the analysis. Any point lying outside the ± 3S limits or seven consecutive points on one side of the mean line $\bar{X}$ will indicate that some systematic error is probably pre-



Quality control chart to monitor accuracy.

sent and corrective action should be taken. Before one set of standard aliquots is exhausted, another set should be prepared, permitting a period of overlap, so that a new quality control chart can be established and it can be seen that comparable accuracy is being obtained with the new standard as compared with the earlier one.

6.2.2. Precision

The precision of the hemoglobin estimation is monitored by making a set of aliquots (a minimum of four) from a single sample. Where venous blood is collected, this can be done by taking extra blood from one individual and putting it into four different bottles which are labelled and distributed randomly amongst other specimens so that the laboratory worker doing the estimations cannot distinguish them from routine specimens. When measurements are made on capillary blood, it may be difficult to get multiple aliquots from one finger prick. In this situation, precision may be monitored by taking a sample of venous blood in the field and then preparing four or more 20 µl aliquots from this, adding each to 4.0 ml of diluent and forwarding them blind as above to the laboratory doing the estimations. From the results on these aliquots, a daily estimate of variability can be obtained and plotted on a control chart for S. The control chart for S is established as follows. In general, if n replicate aliquots for each of k samples chosen to be representative of the range of expected values (e.g., for hemoglobin 8 to 14 g/dl) are sent into the laboratory, the mean population variance ($\bar{S}$) is estimated by:

$$\bar{S} = \sqrt{\frac{1}{k} \cdot \Sigma_i S_i^2}$$

Where S_i^2 is the estimated variance between replicates in the ith sample.

S_i is computed as follows: If X_{i1}, X_{i2}, X_{i3} and X_{i4} are the estimated values of four replicate aliquots of the ith sample, then:

$$S_i^2 = \frac{(X_{i1}^2 + X_{i2}^2 + X_{i3}^2 + X_{i4}^2) - \frac{(X_{i1} + X_{i2} + X_{i3} + X_{i4})^2}{4}}{4}$$

or in general

$$S_i = \sqrt{\frac{\Sigma X_i^2 - \frac{(\Sigma X_i)^2}{n}}{n^*}}$$

where n is the number of replicate aliquots.

A control chart (S chart) can then be constructed as follows:

Central line	$= \bar{S}$
Upper control limit	$= B_4 \cdot \bar{S}$
Lower control limit	$= B_3 \cdot \bar{S}$
Upper warning limit	$= B_4' \cdot \bar{S}$
Lower warning limit	$= B_3' \cdot \bar{S}$

where B_4, B_3, and B_4', and B_3' are given as in Table 4 depending on the value of n (the number of replicates of each specimen)

TABLE 4.
FACTORS FOR CONTROL CHART FOR S

No. of Replicates (n)	B_4	B_3	B_4'	B_3'
4	2.266	0	1.844	0.156
5	2.089	0	1.726	0.274
6	1.970	0.030	1.647	0.353
7	1.882	0.118	1.588	0.412

The following is an example of the construction of a control chart for S for hemoglobin estimation.

Samples of venous blood are taken from a number of subjects — preferably about 40 (although only 20 are shown for the sake of brevity) and the hemoglobin concentration determined on at least four replicate aliquots of each and S^2 calculated (Table 5).

*By international convention n is used rather than n-1 for making routine calculations of S in quality control work.

TABLE 5.
DATA FOR CONTROL CHART FOR S

Subject Specimen Number	Estimated Hb(g/dl) on Replicate Aliquots				
	X_1	X_2	X_3	X_4	S^2
1	8.9	9.0	8.8	8.9	0.0045
2	9.2	9.1	9.3	9.1	0.0069
3	9.4	9.5	9.3	9.3	0.0069
4	9.7	9.6	9.7	9.8	0.0050
5	10.3	10.2	10.4	10.1	0.0125
6	8.6	8.5	8.7	8.5	0.0069
7	11.0	11.1	10.9	11.1	0.0069
8	12.3	12.2	12.3	12.4	0.0050
9	9.2	9.3	9.1	9.0	0.0125
10	11.3	11.1	11.4	11.2	0.0125
11	8.9	8.7	9.0	8.8	0.0125
12	9.7	9.9	9.6	9.6	0.0150
13	10.4	10.2	10.3	10.1	0.0125
14	8.0	8.1	8.2	8.1	0.0050
15	13.4	13.2	13.5	13.3	0.0125
16	10.6	10.4	10.5	10.7	0.0125
17	9.2	9.1	9.0	9.2	0.0069
18	7.8	7.7	7.6	7.8	0.0069
19	12.4	12.6	12.3	12.5	0.0125
20	9.8	9.7	9.6	9.9	0.0125
					$\Sigma S^2 = 0.1889$

$$\bar{S} = \sqrt{\frac{1}{k} \cdot \Sigma_i S_i^2} = \sqrt{\frac{0.1889}{20}} = 0.097$$

The central line of the chart is thus 0.1

The upper control limit	= 2.266	× 0.097	= 0.22
The lower control limit	0	× 0.097	= 0
The upper warning limit	1.844	× 0.097	= 0.18
The lower warning limit	0.156	× 0.097	= 0.02

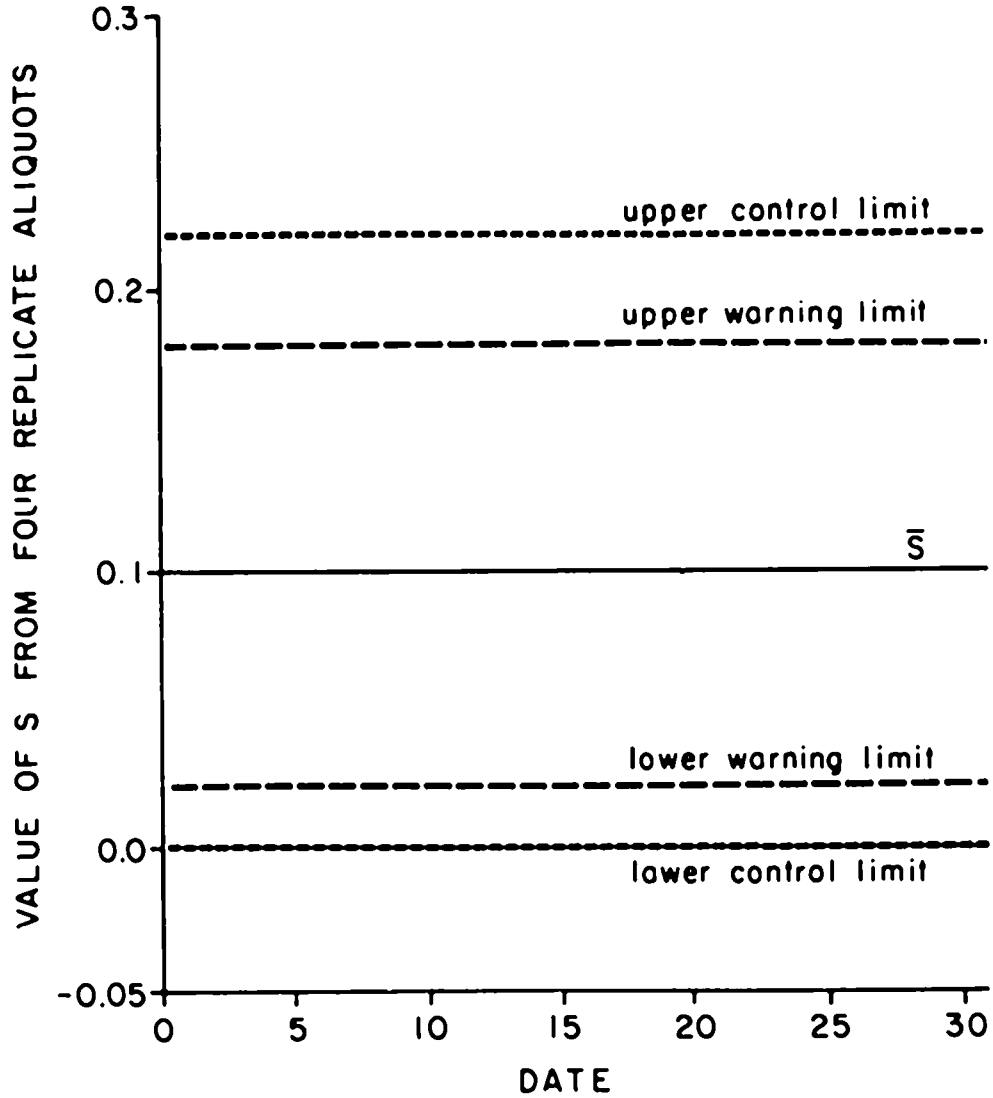
With each batch of specimens, at least one should be divided into four (or more) aliquots and submitted blind to the laboratory worker. The variability (S) of these readings should be determined and the results plotted on the chart (Figure 6). If any estimate of S falls outside the control limits, the method is out of control and the reason must be determined and the measurements obtained in this batch must be discarded. If material is available, the measurements can be repeated when corrective action has been taken. If any points falls between the warning and control limits, or if seven consecutive points fall on one side of the central line, the reasons must be investigated and corrective action taken if necessary.

6.3 Quality control of blood collection.

To maintain some degree of quality control of blood collection, two independent specimens should be obtained from, for example, every fifth or tenth subject sampled by a given worker and these should be submitted blind to the laboratory. This is particularly important when capillary samples are being collected.

Initially, a control chart for each blood collector can be prepared by having the worker take duplicate specimens from about 10 individuals. The variance, S, of these measurements is then calculated and upper and lower warning and control limits are then set at +2S and +3S and −2S and −3S, respectively (see example below). Each day, the difference (d) between duplicate specimens collected by the worker is plotted on that worker's chart corresponding to the date of collection. Inspection of these charts provides assurance that precision of collection continues to be adequate. Because the warning limit is based on approximately 95% confidence limits, a point can be expected to fall be-

FIGURE 6



Control chart for S.

tween the warning and control limits once in 22 occasions. Because of this, no action is called for unless points fall within this range more frequently. If this occurs, it is indicative of a decrease in precision which should be investigated.

Example:
Blood collector #3
Initially, the collector takes duplicate samples from ten individuals and the hemoglobin estimations are made blind (Table 6).

TABLE 6.
DATA FOR CONTROL CHART OF BLOOD COLLECTION.

Subject Number	*Estimated Value*		*Difference Between Duplicates (d)*	d^2
	Specimen 1	*Specimen 2*		
1	8.9	8.9	0.0	0.0
2	9.2	9.1	0.1	0.01
3	10.0	10.2	−0.2	0.04
4	11.5	11.7	−0.2	0.04
5	10.4	10.2	0.2	0.04
6	13.3	13.3	0.0	0.0
7	8.5	8.4	0.1	0.01
8	10.0	9.8	0.2	0.04
9	10.8	10.8	0.0	0.0
10	9.1	8.9	0.2	0.04

$n = 10$ $\Sigma d^2 = 0.22$

$$S = \sqrt{\frac{0.22}{2 \times 10}} = 0.1048$$

A control chart for collection is then constructed:

Central line	=	0	
Upper control limit	=	3S =	0.31
Upper warning limit	=	2S =	0.21
Lower control limit	=	−3S =	−0.31
Lower warning limit	=	−2S =	−0.21

See Figure 7.

6.4. Allocation to placebo and treatment groups

The optimum allocation of subjects to treatment and control (placebo) groups is given by the formula:

$$c = t\sqrt{m}$$

where c is the number of subjects in the control group, t is the number of subjects in each treatment group, and m is the number of treatment groups, excluding the control group.

The random allocation of subjects in these proportions for different numbers of treatment groups can be carried out as shown in Table 7.

TABLE 7.
SCHEMA FOR ALLOCATION OF SUBJECTS TO STUDY GROUPS.

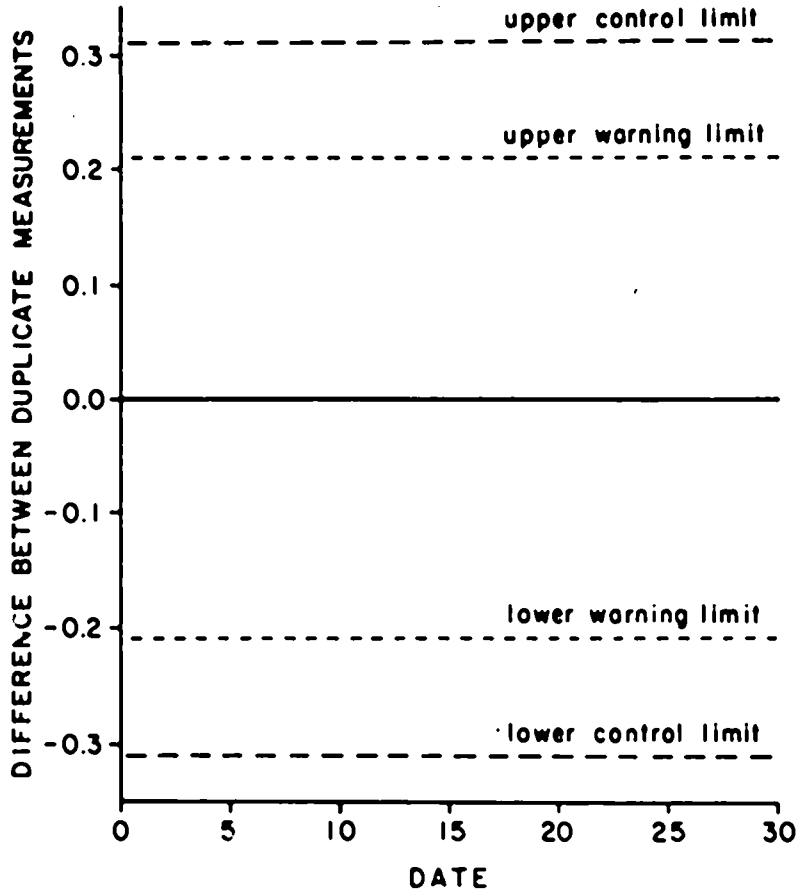
No. of Treatment Groups	*No. of Study Groups (Including Placebo)*	*Cyclical Order of Allocation*	*Suggested Size of Random Permutation Blocks*
1	2	*PA, PA, PA, PA, PA	10
2	3	PAB, PAB, PAB, P	10
3	4	PABC, PABC, PABC, PP	14
4	5	PABCDP, PABCDP, PABCDP	18
5	6	PABCDEP, PABCDEP, PABCDEP, P	22
6	7	PABCDEFP, PABCDEFP, PABCDEFP, P	25

*P denotes placebo and other letters denote the different treatment groups

The blocks of random permutation tables may either be obtained from books of published tables or generated by the lottery procedure.

For example, to generate blocks of ten numbers, the numbers 0, 1, 2, 3, ... 8, 9 are written on identical spheres (e.g., table tennis balls). The

FIGURE 7



Control chart for precision, for a given blood collector, on measurements of duplicate samples.

spheres are then thoroughly mixed in a closed container. The spheres are picked up, one by one, by a blindfolded person and the numbers are noted down by an assistant in the order in which they appear. The numbers should be permanently recorded to facilitate supervisory checks. Blocks of any size can similarly be generated by increasing the number of spheres.

6.5. Determination of sample size for the comparison of two means

Data needed for the estimation of sample size required for comparison of two mean values are:

(i) The difference (Δ) in mean hemoglobin concentration between the two groups considered to be of biological importance.

(ii) The standard deviation (σ) of the hemoglobin concentrations.

(iii) The required confidence coefficient—usually the 95% confidence coefficient is used.

(iv) The probability with which the expected difference, if present, is to be detected. This is often called the 'power' of the test.

Given these, the formula for estimating the sample size in each group (assumed to be of equal size) is:

$$n = \frac{2 \cdot \sigma^2 (Z_\beta + Z_\alpha)^2}{\Delta^2}$$

Where n is the sample size; σ is the standard deviation of Hb concentration, Z_β is the normal deviate corresponding to the 'power' required. (For example, Z_β = 1.645 corresponding to a power of 95%, 1.282 corresponding to a power of 90% and 0.84 corresponding to a power of 80%.) Z_α is the normal deviate corresponding to the confidence coefficient. (For example Z_α = 1.96 corresponding to a 95% confidence coefficient.)

The sample sizes for various values of Δ, σ, 'power' and confidence coefficient are presented in Table 8.

TABLE 8.
ESTIMATED MINIMUM SAMPLE SIZE IN EACH GROUP CORRESPONDING TO DIFFERENT VALUES OF Δ (CHANGE IN HEMOGLOBIN CONCENTRATION) AND σ (STANDARD DEVIATION OF HEMOGLOBIN CONCENTRATION)

					Power = 90%, Confidence Coefficient = 95%						
Δ =	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.2	1.4
σ					*n in each group*						
1.0	525	234	131	84	58	43	33	26	21	15	11
1.2	757	336	189	121	84	62	47	37	30	21	15
					Power = 80%; Confidence Coefficient = 95%						
Δ =	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.2	1.4
σ					*n in each group*						
1.0	392	174	98	63	44	32	24	19	16	11	8
1.2	564	251	141	90	63	46	35	28	23	16	12

6.6. Student's 't' test

When there are two independent samples with mean values $\bar{X}_1$ and $\bar{X}_2$, respectively, and it is desired to test if these two mean values are significantly different or not, the Student's 't' test is used. This test is only applicable under the assumption that the two means come from normally distributed data. The initial assumption is made that the two samples come from the same population, i.e., that there is no significant difference between the two means (this is known as the "null hypothesis"). This is put to the test as follows: a pooled estimate of the population variance σ^2 is obtained from the two samples which is designated as pooled S^2. From this is derived the standard error of the difference of the mean values of the two groups $S_{\bar{X}_1 - \bar{X}_2}$ (see the following example).

The 't' statistic is then calculated by dividing the difference between the two means by the standard error of the difference in the mean values.

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{S_{\bar{X}_1 - \bar{X}_2}}$$

This statistic is then compared with the theoretical tabulated* value of 't' corresponding to the degrees of freedom, to obtain the probability. If this probability is greater than the level of significance (usually 5% or 0.05), it may be concluded that the null hypothesis holds and that there is no significant difference between the two groups in their mean values.

As an example, consider the data in Table 9, which presents the initial hemoglobin concentration of women randomly allocated to receive a placebo (Group 1) or iron (Group 2) in a specified dose.

TABLE 9.

INITIAL Hb CONCENTRATION (g/dl) IN TWO GROUPS OF WOMEN

Group 1	*Group 2*
9.4	9.6
8.4	9.8
9.8	8.4
9.2	9.2
9.5	9.6
8.2	8.4
8.7	8.2
10.8	10.7
10.0	10.2
11.1	10.9
10.2	10.8
11.6	—

It is desired to test whether, after allowing for dropouts, the random allocation has resulted in two similar groups with respect to the initial hemoglobin concentrations. The calculations shown in Table 10 are made.

*In appropriate statistical tables [14]

TABLE 10.
CALCULATIONS FROM DATA IN TABLE 9.

	Group 1		Group 2	
Sum of individual values	ΣX_1	116.9	ΣX_2	105.8
Number of observations	n_1	12	n_2	11
Mean of values	$\bar{X}_1$	9.74	$\bar{X}_2$	9.62
Sum of squares of values	ΣX_1^2	1151.23	ΣX_2^2	1027.34
Sum of values, squared, divided by number in group	$\frac{(\Sigma X_1)^2}{n_1}$	1138.8	$\frac{(\Sigma X_2)^2}{n_2}$	1017.6
Sum of squares of differences of value from respective mean	$\Sigma(X_1-\bar{X}_1)^2$	12.429	$\Sigma(X_2-\bar{X}_2)^2$	9.736
Degrees of freedom	$n_1 - 1$	11	$n_2 - 1$	10

The pooled estimate of the population variance is given by the formula:

$$\text{pooled } S^2 = \frac{\Sigma(X_1-\bar{X}_1)^2 + \Sigma(X_2-\bar{X}_2)^2}{(n_1-1)+(n_2-1)} = \frac{12.429 + 9.736}{11 + 10} = 1.055$$

The standard error of the difference between the mean values of the two groups is given by the formula:

$$S_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{1.055\left(\frac{1}{12} + \frac{1}{11}\right)} = 0.4288$$

$$\text{Then: } t = \frac{\bar{X}_1 - \bar{X}_2}{S_{(\bar{x}_1 - \bar{x}_2)}} = \frac{9.74 - 9.62}{0.4288} = 0.2798$$

$$\text{Degrees of freedom} = (n_1 - 1) + (n_2 - 1) = 11 + 10 = 21$$

From the tables, the probability is greater than 0.7, i.e., there is no significant difference between the two means.

6.7 One-way Analysis of Variance

When the means of more than two independent groups of observations have to be compared, a one-way analysis of variance technique is used.

Consider the data* in Table 11 on initial hemoglobin concentrations (g/dl) of four groups of women randomly allotted to a placebo (Group I) and three different doses of iron (Groups II, Group III and Group IV).

In testing for the differences in mean values of the four groups the hypothesis (called the "null hypothesis") is made that the four groups of women are from the same population. In other words, it is necessary to test that the random allocation has resulted in four groups that have similar initial hemoglobin concentrations. Under this assumption all 79 observations are distributed about a common mean μ and variance σ^2

*These data are taken from an old study and it should be noted that ideally the number of control subjects in the placebo group should have been more to conform with the requirements given in #6.4

TABLE 11.
INITIAL HEMOGLOBIN CONCENTRATIONS OF WOMEN IN A RANDOMIZED TRIAL AND INITIAL CALCULATIONS

	Group I	*Group II*	*Group III*	*Group IV*	
	9.4	9.6	9.0	9.6	
	9.8	9.6	9.8	9.8	
	9.4	9.2	9.4	9.2	
	9.2	9.1	9.2	9.2	
	9.5	9.6	9.4	9.6	
	9.8	9.7	9.8	9.8	
	9.2	9.5	9.8	9.2	
	9.0	9.3	9.0	9.0	
	9.7	11.2	9.6	9.8	
	9.9	10.0	9.7	9.1	
	11.2	10.2	9.1	11.6	
	10.8	10.2	10.0	10.8	
	10.0	10.8	11.4	10.2	
	10.4	10.0	10.4	11.0	
	11.1	10.2	10.8	10.2	
	10.2	10.4	10.0	10.6	
	11.2	11.2	10.2	10.2	
	11.0	10.4	10.8	11.2	
	11.6	11.2	11.8		
	10.8	10.2			
	11.2				
	11.2				
					Total
n:	22	20	19	18	79
ΣX:	225.6	201.6	189.2	180.1	796.5
$\bar{X}$:	10.3	10.1	10.0	10.0	—
ΣX^2:	2327.44	2040.20	1895.22	1812.29	8075.15
$\frac{(\Sigma X)^2}{n}$:	2313.4255	2032.1280	1884.0337	1802.0006	8031.5878
$\Sigma(X-\bar{X})^2$:	14.0145	8.0720	11.1863	10.2894	—

where n is the number of observations
ΣX is the sum of the hemoglobin values
and $\bar{X}$ is the mean hemoglobin value in each group.

The analysis of variance is developed from the fact that three different estimates of the variance (σ^2) can be made from the data given in Table 11. If, indeed, the four groups of women come from the same population, these estimates should not differ significantly from each other.

Estimate I: Total sum of squares.

The total sum of squares of deviations of all 79 observations from the overall mean value is computed as follows:

$$(9.4)^2 + (9.8)^2 + \ldots + (10.2)^2 + (11.2)^2 - \frac{(796.5)^2}{79} = 44.6152$$

where 796.5 is the grand total of all the observations as given in the last column of Table 11. This sum of squares has 78 degrees of freedom, giving a mean square of $\frac{44.6152}{78} = 0.5719$ as the first estimate of σ^2.

Estimate II: Sum of squares within groups.

The second estimate of σ^2 is obtained by computing the sum of squares between the hemoglobin values of women within each group, called the sum of squares within groups. The procedure is as follows: add up the values of $\Sigma(X - \bar{X})^2$, given in the last row of the table, for the four groups, i.e.,

$$14.0145 + 8.0720 + 11.1863 + 10.2894 = 43.5622.$$

This has $(22 + 20 + 19 + 18 - 4) = 75$ degrees of freedom giving the second estimate of σ^2 as

$$\frac{43.5622}{75} = 0.5808.$$

Estimate III: Sum of squares between groups.

The third estimate of σ^2 is obtained as follows: square the sum of the hemoglobin values in each group and divide by the number in the group (the second last line of Table 11). Add the resultant quantities and subtract from this the sum of all the observations squared, divided by the total number of observations

$$2313.4255 + 2032.1280 + 1884.0337 + 1802.006 - \frac{(796.5)^2}{79} = 1.0530.$$

Since there are four groups this sum of squares has three degrees of freedom giving the third estimate of σ^2 as $\frac{1.0530}{3} = 0.3510$.

Once these three estimates are obtained, they can be presented in the form shown in Table 12:

TABLE 12.
ANALYSIS OF VARIANCE

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*Variance Ratio*
Between groups	3	1.0530	0.3510	<1
Within groups	75	43.5622	0.5808	
Total	78	44.6152		

Note that the degrees of freedom, as well as the sum of squares, in the between groups and within groups rows add to the total given in the last row. In the last column of the analysis of variance table, the ratio of the mean square of between groups and within groups is computed. If the means of the four groups differ significantly, then the between groups mean square will be significantly greater than the within group mean square. This is tested by a test called the "F" test or variance ratio test.

An excerpt of the "F" table is shown in Table 13.

TABLE 13.
5% POINTS FOR DISTRIBUTION OF F[14]

n_2 \ n_1	1	2	3	4	5	6	8	12	24
1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	243.9	249.0
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.41	19.45
3	10.13	9.55	9.28	9.12	9.01	8.94	8.84	8.74	8.64
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.77
—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—
40	4.03	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.79
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.70
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.83	1.61
	3.84	2.99	2.60	2.37	2.21	2.10	1.94	1.75	1.52

Note: n_1 stands for the degrees of freedom of the between groups mean square and n_2 stands for the degrees of freedom of the within group mean square in the context of the one-way analysis of variance.

If the variance ratio is less than 1, then there is no need to refer to this table since it means that the variance ratio is not statistically significant. If the variance ratio computed in the analysis of variance table is greater than 1, then the tabulated value of F in the above table should be looked at corresponding to n_1 and n_2. If the tabulated value of F is smaller than the computed value of the variance ratio, then "F" is said to be statistically significant at 5%.

6.8. Two-way analysis of variance

Illustrative data for a two-way analysis of variance and the details of the computations are presented in Tables 14a and b. The two-way analysis of variance is developed as is the one-way analysis of variance, from the fact that we can make four different estimates of the variance (σ^2) from the data given, one estimate corresponding to the variability between treatment groups, one corresponding to the variability between the initial hemoglobin strata, a third estimate corresponding to the variability due to the interaction between treatments and initial hemoglobin strata and a fourth corresponding to the variability left after all the above three are accounted for. These estimates are then presented as in Table 15 and the three estimates of σ^2 (corresponding to between treatment groups, between initial hemoglobin strata and interaction) are individually compared to the fourth estimate by an F test, as was done in the case of the one-way analysis of variance in #6.7.

The analysis presented here is applicable only when the number of observations in all the cells of

TABLE 14a.
CHANGE IN HEMOGLOBIN CONCENTRATION (δ) AFTER THERAPY IN THREE RANDOMIZED GROUPS OF WOMEN

INITIAL Hb (gm %)		GROUP I	GROUP II	GROUP III	
8 – 10		-0.2 +0.1 0.0 -0.7 +0.3 +0.6 -0.9	+1.2 +1.4 +0.8 +2.6 +1.4 +2.0 +2.1	+1.0 +1.5 +3.4 +1.9 +3.0 +2.1 +1.4	
	Cell Total*	C_1 = -0.8	C_2 = +11.5	C_3 = +14.3	$(C_1+C_2+C_3)$ = +25.0
	No. of Observations**	n_1 = 7	n_2 = 7	n_3 = 7	$n_1+n_2+n_3$ = 21
>10		+1.0 +0.1 -0.2 -0.2 +0.4 +0.8 +0.1 -1.4 +1.0 +0.2 -1.3 +1.2 -1.4 -0.5	0.0 +1.2 +1.4 +1.2 +1.1 -0.1 -0.6 +0.7 +1.4 +0.5 +1.8 +0.6 +1.3 +0.7	+1.6 -0.8 +1.6 +2.2 +1.4 +0.6 +1.0 +0.7 +2.3 +2.9 +0.6 +1.2 +1.7 +0.9	
	Cell Total	C_4 = -0.1	C_5 = +11.2	C_6 = +17.9	$(C_4+C_5+C_6)$ = +29.0
	No. of Observations	n_4 = 14	n_5 = 14	n_6 = 14	$n_4+n_5+n_6$ = 42
	Group Total	C_1+C_4 = -0.9	C_2+C_5 = +22.7	C_3+C_6 = +32.2	GRAND TOTAL $T = C_1+C_2+C_3+C_4+C_5+C_6$ = +54.0
	No. of Observations	n_1+n_4 = 21	n_2+n_5 = 21	n_3+n_6 = 21	$n = n_1+n_2+n_3+n_4+n_5+n_6$ = 63

* Each one of the 6 subclasses in the table is called a "Cell". For example, the subclass "Initial Hb 8-10 gm % in GROUP I" is one cell – called C_1

** n_i is the number of observations in the ith cell

TABLE 14b.
COMPUTATIONS REQUIRED FOR THE ANALYSIS

(i) Correction Factor (C.F.)

$$\frac{(T)^2}{N} = \frac{(54)^2}{63} = 46.2857$$

(ii) Total sum of squares (= Total SS)

(Square of each observation) − C.F.

$$= (-0.2)^2 + (0.1)^2 + (-0.7)^2 + \ldots + (0.9)^2 - C.F. = 115.76 - 46.2857 = 69.4743$$

(iii) Between Cells SS

$$= \frac{C_1^2}{n_1} + \frac{C_2^2}{n_2} + \frac{C_3^2}{n_3} + \frac{C_4^2}{n_4} + \frac{C_5^2}{n_5} + \frac{C_6^2}{n_6} - C.F.$$

$$= \frac{(-0.8)^2}{7} + \frac{(11.9)^2}{7} + \frac{(14.3)^2}{7} + \frac{(-0.1)^2}{14} + \frac{(22.7)^2}{14} + \frac{(32.2)^2}{14} - C.F. = 33.75$$

(iv) Between Groups SS

$$= \frac{(\text{Total of Group I})^2}{n_1 + n_4} + \frac{(\text{Total of Group II})^2}{n_2 + n_5} + \frac{(\text{Total of Group III})^2}{n_3 + n_6} - C.F.$$

$$= \frac{(-0.9)^2}{21} + \frac{(22.7)^2}{21} + \frac{(32.2)^2}{21} - C.F. = 27.66$$

(v) Between Initial Hb Strata SS

$$= \frac{(C_1 + C_2 + C_3)^2}{n_1 + n_2 + n_3} + \frac{(C_4 + C_5 + C_6)^2}{n_4 + n_5 + n_6} - C.F.$$

$$= \frac{(25.0)^2}{21} + \frac{(29.0)^2}{42} - C.F. = 3.50$$

(vi) Interaction: Treatment Group X Initial Hb Strata

Between Cell SS − Between Group SS − Between Initial Hb Strata SS

$$= 33.75 - 27.66 - 3.50 = 2.59$$

(vii) Within Cells SS

= Total SS − Between Cell SS

$$= 69.4743 - 33.75 = 35.7243$$

the table are either equal, or the observations in the subclasses within treatment groups are proportional as in Table 14. That is, the proportion between the number of observations in the 8 to 10 g and >10 g initial Hb strata within each treatment group is the same (1:2 in the present example). If the observations in the sub-classes are unequal and disproportionate, then the analysis is a little more

TABLE 15.
ANALYSIS OF VARIANCE

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Square*	*Variance Ratio*	*Probability*
Between Groups	2	27.66	13.83	$\frac{13.83}{0.63} = 21.95$	<0.001
Between Initial Hb Strata	1	3.50	3.50	$\frac{3.50}{0.63} = 5.56$	<0.05
Interaction: Groups X Initial Hb Strata	2	2.59	1.29	$\frac{1.29}{0.63} = 2.05$	NS
Within Cells	57	35.7243	0.63	—	

complex and it would be advisable to consult a competent statistician to carry out and interpret the analysis.

6.9. Analysis of covariance

The analysis of covariance combines features of regression technique and analysis of variance. The analysis of covariance has, among others, two important uses relevant to randomized experiments of the type considered in this publication, namely (i) to increase precision in the comparison of hemoglobin concentrations after therapy in the different groups by removing the effects of differences in the initial hemoglobin concentrations and (ii) to compare the regressions in each of the groups to see if the slopes of the regression lines differ or not. If the slopes do not differ, a combined estimate of the slope can be made for all the groups. Otherwise, separate regression lines have to be used for each group to remove the effect of initial hemoglobin concentrations.

The rationale of the technique is as follows:

Let Y_{ij} represent the hemoglobin concentration after therapy in subject number j in group i.
A linear model is set up

$$Y_{ij} = \mu_i + \beta (X_{ij} - \bar{X}_{..}) + \epsilon_{ij}$$

where μ_i represents the effect of the treatment; X_{ij} represents the initial hemoglobin concentration of subject j in group i, β is the regression coefficient of Y on X, $\bar{X}_{..}$ is the mean initial hemoglobin concentration of all subjects in all groups and ϵ_{ij} is a random error term.

The observed mean for the ith group then is:

$$\bar{Y}_i = \mu_i + \beta (\bar{X}_{i.} - \bar{X}_{..}) + \bar{\epsilon}_i$$

It follows from the above, therefore, that $\bar{Y}_i - \beta (\bar{X}_{i.} - \bar{X}_{..})$ is an estimate of μ_i, the term $[-\beta (\bar{X}_{i.} - \bar{X}_{..})]$ being the adjustment introduced by the covariance analysis. The value of β in the above model is estimated by the formula:

$$b = \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{\Sigma(X - \bar{X})^2}$$

(The symbol b is used since it is an estimate of β)

An example of the use of analysis of covariance in analyzing the results of a supplementation trial follows. Table 16a presents the initial and final hemoglobin concentrations and basic computations for women randomized to placebo (Group I) and two different supplementation strategies (Groups II and III).*

In comparing a number of treatment groups, the question is whether the regression lines in the different groups are the same, that is, are the slopes and the elevations of the regression lines the same, or are they different? The computations required for comparing the slopes and the elevations for the data given in Table 16a are presented in Table 16b.

---

*See footnote page 29

TABLE 16a.
DATA FROM A SUPPLEMENTATION TRIAL

Initial (X) And Final (Y) Hemoglobin Concentrations In Three Randomized Groups Of Women

GROUP I		GROUP II		GROUP III	
X	Y	X	Y	X	Y
9.4	9.2	9.6	8.9	9.8	10.8
8.4	8.5	9.8	12.8	8.9	10.4
9.8	9.8	9.2	10.4	8.0	11.4
9.2	8.5	8.4	9.8	9.7	11.6
9.5	9.8	9.2	9.1	9.9	12.9
8.2	7.2	9.6	12.2	9.2	12.7
8.7	11.0	9.6	11.0	9.2	11.9
9.0	9.6	9.8	11.8	9.8	12.5
9.7	8.8	8.4	11.6	8.4	10.5
9.9	10.5	9.8	11.8	9.7	11.1
9.0	8.2	9.1	11.2	9.5	8.6
9.7	10.9	9.9	12.2	9.4	11.1
8.0	9.0	9.6	10.4	9.3	9.7
8.6	8.0	9.4	9.7	9.7	13.1
11.2	9.6	9.7	10.9	9.3	9.1
10.8	11.8	9.3	10.6	9.0	12.7
10.0	11.6	7.8	12.3	8.6	9.1
11.1	11.2	11.6	11.6	10.4	12.0
10.2	10.2	10.8	10.4	10.6	12.0
11.0	10.8	10.2	11.4	10.8	10.0
11.6	9.8	10.2	11.6	10.0	11.6
10.8	10.6	10.6	11.8	10.8	13.0
11.2	11.6	11.2	11.0	10.6	12.0
11.2	12.2	10.8	11.9	10.2	10.8
10.5	9.2	11.0	10.9	10.4	11.4
10.0	10.8	11.8	11.2	12.4	13.1
10.3	10.4	11.6	12.3	10.8	13.1
10.8	9.4	11.4	12.8	10.0	12.9
11.0	11.0	11.9	11.8	11.4	12.0
11.2	12.2	11.0	11.5	12.3	11.8
11.6	11.8	10.3	12.1	10.2	11.4
10.5	9.2	10.7	11.3	11.0	11.7
10.8	12.1	11.9	13.2	12.2	13.1
10.0	8.6	11.9	13.2	10.0	12.5
11.4	10.9	10.6	11.3	10.2	11.7
11.4	10.4	12.0	11.9	12.7	12.4
11.1	10.4	10.6	12.5	11.9	14.6
10.3	9.9	10.7	13.4	11.0	12.4
				11.5	11.8
				10.2	12.6
				11.6	13.4
				12.6	13.9
				11.8	12.2
				11.8	11.8
				10.4	10.7
				12.8	13.4
				11.2	12.3
				11.2	12.9
				10.1	12.5
				10.3	10.9
				11.8	10.7
				10.7	10.4
				11.2	10.4

basic computation required for each group

1 Degrees of freedom (n-1)	37	37	52
2 ΣX, ΣY	387.10 ; 384.70	391.00 ; 435.80	556.50 , 624.60
3 ΣX^2, ΣY^2	3981.4300, 3953.8700	4065.8600; 5039.5800	5912.0300 , 7441.9000
4 ΣXY	3950.1100	4502.0300	6594.5500
5 $\Sigma(X-\bar{X})(Y-\bar{Y})$ $=\Sigma XY-(\Sigma X\,\Sigma Y)/n$	31.2318	17.8774	36.2500
6 $\Sigma(X-\bar{X})^2$ $=\Sigma X^2-(\Sigma X)^2/n$	38.1034	42.6758	68.7800
7 $\Sigma(Y-\bar{Y})^2$ $=\Sigma Y^2-(\Sigma Y)^2/n$	59.2887	41.6421	81.0479

TABLE 16b.
ANALYSIS OF COVARIANCE: COMPARISON OF REGRESSION LINES IN THE THREE GROUPS

Row No.	Source of Variation	Degrees of Freedom	$\Sigma(X-\bar{X})^2$	$\Sigma(X-\bar{X})(Y-\bar{Y})$	$\Sigma(Y-\bar{Y})^2$	Regression Coefficient	Deviations from Regression: Degrees of Freedom	Deviations from Regression: Sum of Squares	Deviations from Regression: Mean Squares
	Within Groups								
1	Group I	37	38.1034	31.2318	59.2887	0.820	36	33.6893	0.9358
2	Group II	37	42.6758	17.8774	41.6421	0.419	36	34.1330	0.9487
3	Group III	52	68.7800	36.2500	81.0479	0.527	51	61.9426	1.2146
4		Sum of Deviations from Regression					123	129.7849	1.0552
5	Pooled within groups (W)	126	149.5592	85.3592	181.9787	0.571	125	133.2609	1.0661
6		Comparison between the slopes F = .7380 = 1.0552 ÷ 1.66, P > 0.0 (Slopes are not significantly different)					2	3.4760	1.7380
7	Between groups (B)	2	2.34.2	0.7073	64.8903	-	-	-	-
8	(W) + (B)	28	5.9004	96.0445	246.8690	-	127	186.1136	1.4653
9		Differences between adjusted means F = 26.4264 = .066 ÷ 24.79, P < 0.001 (Means are significantly different)					2	52.8527	26.4264
10	Pooled within groups (W) (groups II and III only)	89	111.4558	54.1274	122.6900	0.4856	88	96.4036	1.0955
11	Between groups (B_1) (groups II and III only)		0.9809	4.750	22.68	-	-	-	-
12	(W_1) + (B_1)	90	112.4367	55.6024	124.9048	-	89	97.4102	1.0943
13		Difference between adjusted means (groups II and III only) F = 1.0066 = 1.0955 ÷ 0.989 (Deviations are not significantly different)					1	1.0066	1.0066

In rows 1, 2 and 3, of Table 16b, the regression coefficients are estimated separately for each group. The right hand extreme of the table titled "Deviations from Regression," presents the residual sum of squares and the residual mean square of the final hemoglobin concentrations in each group after the regression line is fitted.

The residual sum of squares is obtained by using the formula.

$$\Sigma(Y-\bar{Y})^2 - b^2 \cdot \Sigma(X-\bar{X})^2$$

where b is the estimate of the regression coefficient.

The residual mean square is obtained by dividing the residual sum of squares by the appropriate degrees of freedom (given under "Deviations from Regression"). In row 4, the residual mean square is calculated after fitting the three separate regressions to the three groups. In row 5, a common slope for the regression lines is estimated for the three groups by adding up the sums of squares and products from lines 1, 2 and 3 and the corresponding residual mean square estimated. If the residual mean square corresponding to a common slope is *not* significantly higher than that of the residual mean square obtained after fitting three separate slopes to the three groups, it can be concluded that the slopes do not differ significantly. In the example the F test (row 6) is not significant.

The next step is to compare the elevations of the regression lines for significant differences. This test, mathematically, is identical with the test for the equality of the mean post-therapy hemoglobin concentrations after adjustment for differences in the initial hemoglobin concentrations, as can be seen from the following. The estimate of the elevations (denoted usually by α_I, α_{II} and α_{III}) for the three regression lines in the example, are given by:

$$\alpha_I = \bar{Y}_I - b\,\bar{X}_I$$

$$\alpha_{II} = \bar{Y}_{II} - b\,\bar{X}_{II}$$

$$\alpha_{III} = \bar{Y}_{III} - b\,\bar{X}_{III}$$

where $\bar{Y}_I$, $\bar{Y}_{II}$, $\bar{Y}_{III}$ stand for the mean hemoglobin after therapy in the first, second and third groups, and $\bar{X}_I$, $\bar{X}_{II}$, $\bar{X}_{III}$, stand for the mean initial hemoglobin in the three groups, respectively. It is obvious that these represent the adjusted mean of the post-therapy hemoglobin concentrations. For this reason, row 9 of Table 16b is written as "differences between adjusted means."

The "between groups" sums of squares and products in row 7 are computed as follows:

Between groups $\Sigma(X - \bar{X})^2$

$$= \frac{(\Sigma X_I)^2}{n_I} + \frac{(\Sigma X_{II})^2}{n_{II}} + \frac{(\Sigma X_{III})^2}{n_{III}} - \frac{(\Sigma X_I + \Sigma X_{II} + \Sigma X_{III})^2}{n_I + n_{II} + n_{III}}$$

Between groups $\Sigma(X - \bar{X})(Y - \bar{Y})$

$$= \frac{(\Sigma X_I \cdot \Sigma Y_I)}{n_I} + \frac{(\Sigma X_{II} \cdot \Sigma Y_{II})}{n_{II}} + \frac{(\Sigma X_{III} \cdot \Sigma Y_{III})}{n_{III}} - \frac{(\Sigma X_I + \Sigma X_{II} + \Sigma X_{III})(\Sigma Y_I + \Sigma Y_{II} + \Sigma Y_{III})}{n + n_{II} + n_{III}}$$

Between groups $\Sigma(Y - \bar{Y})^2$

$$= \frac{(\Sigma Y_I)^2}{n_I} + \frac{(\Sigma Y_{II})^2}{n_{II}} + \frac{(\Sigma Y_{III})^2}{n_{III}} - \frac{(\Sigma Y_I + \Sigma Y_{II} + \Sigma Y_{III})^2}{n_I + n_{II} + n_{III}}$$

where ΣX_I, ΣX_{II}, ΣX_{III} stand for the sum of all the X values and ΣY_I, ΣY_{II}, ΣY_{III}, stand for the sum of all the Y values in groups I, II, and III, respectively, and n_I, n_{II} and n_{III} stand for the total numbers in these groups

The computations in rows 8 and 9 of the table are self-explanatory. The tests in row 6 and row 9 taken together show that the slopes of the regression lines are not significantly different, but their elevations are. The computations in rows 10, 11, 12 and 13 test if the elevations of the regression lines in the case of Group II and Group III are different or not. The conclusion that emerges from the analysis is that the regression lines of the two treatment groups do not differ significantly either in slope (row 6) or in elevation (row 12) and that the regression line of the placebo group has a similar slope (row 6), but a significantly different elevation (row 9).

It is important to note a point of detail in the analysis here. In row 6 of Table 16b, the F ratio is higher than 1, but is not significant. Whenever the F ratio exceeds 2, even if it is not statistically significant, it is safer to treat the slopes as different and examine the estimated slopes in the individual groups to see if the placebo group differs from the rest. If it appears so, then the slopes in the treatment groups, excluding the placebo group, can be compared by an analysis exactly similar to that done in rows 1 to 6.

TABLE 16c.
ESTIMATED REGRESSION LINES AND TREATMENT EFFECT

Estimated Regression Lines

General form $Y = \hat{\alpha} + \hat{\beta} X$

where $\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$

and $\hat{\beta}$ = Regression Coefficient

(i) Placebo group $Y = 4.30 + 0.571 X$

(ii) Treatment groups $Y = 5.71 + 0.571 X$

(Common lines for Groups II & III)

Estimated Post-therapy Hemoglobin Concentrations and Estimated Effect of Treatment at Different Initial Hemoglobin Concentrations

Initial Hb (gm %) X	Estimated Post-therapy Hb (gm %) Y		Estimated effect of Treatment (gm %)
	Placebo	Treatment	
8	8.87	10.28	1.41 *
9	9.44	10.85	1.41
10	10.01	11.42	1.41
11	10.58	11.99	1.41
12	11.15	12.56	1.41

* The effect is constant at all levels of initial Hb since the regression lines of Placebo and Treatment groups are parallel. If they were non-parallel, the treatment effect would be different at different hemoglobin concentrations

Table 16c gives the estimated regression lines for the placebo and the treatment groups. For the two treatment groups, a common regression line is given and a separate regression line is given for the placebo group. From these regression lines, the expected post-therapy hemoglobin concentrations corresponding to any given initial hemoglobin concentration can be estimated by making X equal to the initial hemoglobin concentration. The difference between the estimated post-therapy hemoglobin concentrations in the placebo and treatment groups gives the estimated increase in hemoglobin concentration due to treatment. Since, in this case, the regression lines are parallel (i.e., the slopes are the same), the estimated increase in hemoglobin concentration due to treatment is constant at all initial levels of hemoglobin. If, however, the regression lines are non-parallel, the effect of treatment will be different at different levels of initial hemoglobin concentration and it is desirable to report the estimated effects of treatment at chosen initial levels of hemoglobin separately. Although this example is taken from a supplementation trial, it is more often found that the regression lines are non-parallel and the treatment effect is usually greater at lower initial hemoglobin concentrations.

REFERENCES

1. Guidelines for the Eradication or Iron Deficiency Anemia. A report of the International Nutritional Anemia Consultative Group (INACG, 1978)
2. Nutritional Anemias, World Health Organization Technical Report Series, No 503, WHO Geneva, 1982
3. Garn S M, Smith N J, Clark D C Race differences in hemoglobin levels Ecol Food Nutr 3 299, 1974
4. Garn S M, Ryan A S, Owen G M, Abraham S L Income matched black-white hemoglobin differences after correction for low transferrin saturations Amer J Clin Nutr 34 1645, 1981
5. Williams D M Racial differences of hemoglobin concentration measurements of iron, copper and zinc Amer J Clin Nutr 34 1694, 1981
6. Iron deficiency in Infancy and Childhood A report of the International Nutritional Anemia Consultative Group (INACG, 1979)
7. Pootrakul P, Skikne B S, Cook J D The use of capillary blood for the measurement of circulating ferritin Amer J Clin Nutr 37 307, 1983
8. Piomelli S, Brickman A, Carlos E Rapid diagnosis of iron deficiency by measurement of free erythrocyte porphyrins and hemoglobin the FEP hemoglobin ratio Pediatrics 57 136, 1976
9. International Committee for Standardization in Haematology: Recommendations for haemoglobinometry in human blood Brit J Haematol (Suppl) 13 71, 1967
10. Bothwell, T H, Charlton R W, Cook J D, Finch C A Iron metabolism in man Blackwell, Oxford, 1979
11. International Committee for Standardization in Haematology (Iron Panel) Brit J Haematol 38 291, 1978
12. Labbe R F, Finch C A Smith N J, Dosin R N, Sood S K, Madan N, Erythrocyte protoporphyrin heme ratio in the assessment of iron status Clin Chem 24 87, 1979
13. Laynsse, M, Roche M The relationship between anemia and hookworm infection Am J Hyg 79 279, 1964
14. Fisher, R A, Yates, F Statistical Tables Oliver and Boyd, London, 1957
15. Sood, S K, Ramachandran, K, Mathur M, Gupta, K, Ramalingaswamy V, Swarnabai, C, Ponniah, J., Mathan, V I, and Baker S J WHO sponsored collaborative studies on nutritional anaemia in India. Quart. J Med 44 241, 1975

Other task force reports published by the International Nutritional Anemia Consultative Group are:

Guidelines for the Eradication of Iron Deficiency Anemia
Iron Deficiency in Infancy and Childhood
Iron Deficiency in Women
The Effects of Cereals and Legumes on Iron Availability

These reports are available from:
The Nutrition Foundation
888 Seventeenth Street, N.W
Washington, D.C. 20006
U.S.A