A review of methods to detect cases of severely malnourished children in the community for their admission into community-based therapeutic care programs

Mark Myatt, Tanya Khara, and Steve Collins

Abstract

Background. *The complexity and cost of measuring weight-for-height make it unsuitable for use by community-based volunteers. This has led community therapeutic care programs to adopt a two-stage screening and admission procedure in which mid-upper-arm circumference (MUAC) is used for referral and weight-for-height is used for admission. Such a procedure results in many individuals being referred for care on the basis of MUAC but subsequently being refused treatment because they do not meet the weight-for-height admission criterion. This "problem of rejected referrals" has proved to be a major barrier to program uptake.*

Objective. *To systematically review methods to detect cases of severely malnourished children in the community for their admission into community-based therapeutic care programs.*

Methods. *Clinical and anthropometric methods for case detection of severely malnourished children in the community were reviewed with regard to their ability to reflect both mortality risk and nutritional status.*

Results. *MUAC, with the addition of the presence of bipedal edema, was found to be the indicator best suited to screening and case detection of malnutrition in the community. The case definition "MUAC < 110 mm OR the presence of bipedal edema," with MUAC measured by a color-banded strap, is suitable for screening and case detection of malnutrition in the community for children aged between 6 and 59 months. Monitoring and discharge criteria were also reviewed.*

Conclusions. *There is no compelling evidence to support a move away from using weight in combination with clinical criteria for monitoring and discharge.*

Mark Myatt is affiliated with the Institute of Ophthalmology in London; Tanya Khara and Steve Collins are affiliated with Valid International, Oxford, UK.

Please address queries to the corresponding author: Mark Myatt, Institute of Ophthalmology, 11-43 Bath St., London EC1V 9EL, UK; e-mail: mark@brixtonhealth.com.

Key words: Anthropometry, child mortality, community-based management, mid-upper-arm circumference, severe childhood malnutrition

Introduction

Case detection at the community level and the definition of appropriate referral and admission criteria are important factors in achieving adequate levels of coverage for the treatment of severe malnutrition. These considerations have not, until recently, received much attention, because the delivery of services to the severely malnourished has been dominated by intensive treatment delivered in high-dependency inpatient units at high cost to both the provider (e.g., staffing, infrastructure) and the patient and family (e.g., risk of nosocomial infection, loss of carer for siblings, and loss of labor to household). These high costs lead to a scarcity of provision and are barriers to accessing care that limit program coverage [1, 2].

A new model of delivering care has been proposed, called community-based therapeutic care (CTC), that is designed to address the limitations of inpatient care [3]. CTC programs use decentralized networks of outpatient treatment sites (usually located at existing primary health-care facilities), small inpatient units (usually located in existing local hospital facilities), and large numbers of community-based volunteers to provide case detection and some follow-up of patients in their home environments. Patients with severe malnutrition, with good appetite, and without medical complications are treated in an outpatient therapeutic program (OTP) that provides ready-to-use therapeutic food (RUTF) and medicines to treat simple medical conditions. The food and medicines are taken at home, and the patient attends an OTP site weekly or fortnightly for monitoring and resupply. Severely malnourished persons with medical complications and/or anorexia are treated in an inpatient stabilization center (SC) where they receive standard World Health Organization (WHO)-recommended initial care until they have enough appetite and are well enough to continue with outpatient care [4]. CTC programs have treated more than 9,000 severely malnourished children in Ethiopia, Malawi, and Sudan, meeting Sphere Project targets for clinical outcomes and achieving coverage of over 70% in most cases [5]. The CTC delivery model was conceived, developed, and implemented in complex emergency contexts. There are, however, no compelling technical reasons why the CTC model cannot be implemented in developmental settings. Experience of implementing CTC in transitional and developmental contexts is currently being acquired in Bangladesh, Ethiopia, Malawi, and Zambia.

The WHO manual on the treatment of severe malnutrition recommends that children who have a weight-for-height z-score below –3.00 or a weight below 70% of the median weight-for-height (W/H) according to the National Center for Health Statistics (NCHS) reference population median, or who have bipedal edema, be referred for inpatient treatment [4]. This case definition was devised for use in clinical settings by clinical staff and has proved problematic when used in CTC programs. The complexity and cost of the W/H indicator make it unsuitable for use by community-based volunteers. The use of a two-stage referral and admission system, in which referral is based on mid-upper-arm circumference (MUAC) measured in the community by community-based volunteers, and admission is based on W/H measured at the treatment site by program staff, has proved to be a barrier to accessing care. The use of an adequately sensitive MUAC threshold (i.e., a MUAC threshold likely to identify all or almost all persons meeting the W/H-based admission criteria) results in many patients being referred for care who are then refused treatment because they do not meet the W/H-based admission criteria [6].

Operational research undertaken within CTC programs has found that as a result of this *problem of rejected referrals,* carers of referred children become unwilling to bring their children for admission into the program even when the child's condition deteriorates, carers of rejected children actively disparage the program, local leaders become disillusioned with the program, and the levels of staff and volunteer morale and performance fall [6–9]. In some programs the problem of rejected referrals was solved by moving toward a unified MUAC-based referral and admission criterion [9]. In other situations, where there was institutional resistance to the adoption of a unified MUAC-based referral and admission system, the problem of rejected referrals was solved by instituting a system of incentive payments for carers of referred children [10].

Referral of large numbers of children to treatment sites for second-stage screening by a two-stage system also tends to lead to crowding and long waits at treatment sites and the diversion of often scarce resources away from treatment and carer education toward crowd-control and second-stage-screening activities. Long waits at treatment centers have a negative impact upon the community's perception of programs, and this has a negative impact upon program coverage [6, 11]. Crowding and waiting times could be considerably reduced by the use of a unified (i.e., single-stage) referral and admission system.

Operational research undertaken within CTC programs in developmental settings has found that health workers and carers tend to be confused by the difference between classifications based on weight-for-age (W/A), weight-for-height (W/H), and height-for-age (H/A) in situations in which growth-monitoring programs using W/A or community nutrition programs using H/A are operating. This confusion gives rise to a problem of inappropriate, and thus rejected, referrals, leading to problems with program acceptance and integration with existing health-care providers [12, 13].

It is now clear that the implementation of community-based treatment strategies for severe malnutrition in emergency and developmental contexts will require a reassessment of case-detection methods for severe malnutrition. This report presents a review of the options available for case detection of severely malnourished children in the community suitable for use in programs that follow the CTC model of care delivery.

Selecting an appropriate indicator

Conceptual and methodologic framework

The defining characteristics of an appropriate case-detection method depend upon the context in which case detection is taking place. A failure to account for context may lead to inappropriate case-detection methods being adopted and controversy regarding the appropriateness of adopted methods. Sackett and Holland [14] provide a general, and generally accepted, framework for assessing the appropriateness of case-detection methods in different contexts by scoring the relative importance of a set of properties that may be used to typify all case-detection methods:

» Simplicity: the method can be easily administered by nonclinicians;
» Acceptability: the method is acceptable to the subject and others;
» Cost: the overall cost of the method;
» Precision: the degree of reproducibility among independent measurements of the same true value (also known as *reliability*);
» Accuracy: the proximity of a measurement to its true value;
» Sensitivity: the proportion of diseased subjects who test positive;
» Specificity: the proportion of healthy subjects who test negative;

» Predictive value: the probability that a person with a positive test has the disease or that a person with a negative test does not have the disease.

Sackett and Holland identify four distinct contexts in which case-detection methods are applied: epidemiologic surveys and surveillance, case detection in the community (screening), case-finding in clinical contexts, and diagnosis in clinical contexts.

Beaton and Bengoa [15] recommend that indicators suitable for screening and case detection of malnutrition in the community should, in addition to the properties identified by Sackett and Holland [14], allow for *completeness of coverage* and be both *objective* and *quantitative*. Coverage in this context refers to the coverage of case-detection activities rather than the coverage of the treatment program. This has both a spatial and a temporal component. Completeness of coverage implies that all persons at risk are routinely and repeatedly screened. Coverage of a case-detection method may therefore be seen as a product of simplicity, acceptability, and cost, as well as of factors relating to program organization, rather than as a separate property. In situations of relative resource scarcity, completeness of coverage can only be achieved by simple, acceptable, and low-cost case-detection methods.

Jelliffe and Jelliffe [16] recommend that indicators suitable for detecting cases of malnutrition in early childhood should, in addition to having the properties identified above, be reasonably independent of precise knowledge of the subject's age, since this is often difficult to ascertain accurately in the contexts in which programs treating severe malnutrition are required.

Table 1 reproduces the original analysis of Sackett and Holland [14], modified to include the properties identified by Beaton and Bengoa [15] and Jelliffe and Jelliffe [16].

An important *operational* consideration is who will apply the case-detection method. This report assumes that case-detection methods will be applied by minimally trained community-based volunteers with limited schooling and low levels of numeracy and literacy. For this reason, the relative importance of the simplicity of application has been increased from "moderate," as suggested in the original analysis of Sackett and Holland [14], to "crucial" in **table 1**. The meaning of this property is also changed from the original "easily administered by nonclinicians" to "capable of being administered by minimally trained community-based volunteers with limited schooling and low levels of numeracy and literacy."

The original Sackett and Holland [14] framework places more emphasis on sensitivity (deemed "crucial" in their original framework) than on specificity (deemed "moderate" in their original framework). This lack of emphasis on specificity may be better suited to situations in which *suspected cases* detected by screening and case detection in the community are then confirmed by more precise, accurate, and specific methods in a clinical context (i.e., using methods that meet the requirements that Sackett and Holland [14] specify for case-finding in clinical contexts). In such situations, screening and case-finding in the community refers to screening for referral into a second-stage screen that decides admission rather than screening for

TABLE 1. Relative importance of key properties of case-detection methods in different contexts[a]

Property	Context			
	Epidemiologic survey/surveillance	Screening and case detection in the community	Case-finding in clinical contexts	Diagnosis in clinical contexts
Simplicity	++++	++++	–	–
Acceptability	++++	+++	+	–
Cost	++++	++	–	–
Objectivity	++++	++++	++++	++++
Quantitativeness	++++	++++	–	–
Independence of age	++++	++++	–	–
Precision (reliability)	+ (individual) ++++ (group)	++	++++	++++
Accuracy	+ (individual) ++++ (group)	++	++++	++++
Sensitivity	+	++	+++	+++
Specificity	+	++++	++++	++++
Predictive value	+	++	++++	++++

a. Scoring of importance: – irrelevant, + minor, ++ moderate, +++ major, ++++ crucial. The table reproduces the original analysis of Sackett and Holland [14], modified to include the properties identified by Beaton and Bengoa [15] and Jelliffe and Jelliffe [16].

admission. This report concentrates on case-detection methods that unify referral and admission and allow screening staff to refer children for admission rather than for further screening, because such a procedure avoids the problem of rejected referrals. In a unified referral and admissions system, case-detection methods should be specific as well as sensitive, and the relative importance of these properties will differ from those originally specified by Sackett and Holland [14]. With a case-detection method based around (for example) a threshold value of an anthropometric indicator of nutritional status, a large proportion of deaths in untreated individuals (50% or more) should occur in children below the case-defining threshold. Deaths in children below the case-defining threshold are likely to be related to nutritional status and to respond to dietary treatment. Case-detection methods should, therefore, be highly specific, and a good case-detection method will have reasonable levels of sensitivity at high levels of specificity. For this reason, the relative importance of sensitivity and specificity presented in **table 1** has been reversed from that presented in the original analysis of Sackett and Holland [14].

Habicht [17] reviews the relative importance of the properties of case-detection methods in the contexts of screening and surveillance of nutritional status. In this analysis, the relative costs of misdiagnosis, financial and other, are proposed as an additional property to be considered when selecting a case-detection method. Under situations of scarcity of capacity, this consideration favors the adoption of methods that are designed to match capacity to treat rather than the need to treat. Such methods will usually have high specificity but low sensitivity. A consequence of matching capacity to treat rather than need to treat is that the case-detection method will select only the most extreme cases. This results in a case-detection method that excludes the opportunities offered by early detection and consequent early treatment and resolution, which further exacerbates problems associated with scarcity. The analysis of Habicht [17] seems, therefore, best suited to delivery models that can be characterized by extreme scarcity of capacity relative to need and in which a false positive misdiagnosis may have negative consequences for the subject and the family as well as high financial cost to the provider. It may not be well suited to alternative models of delivery, such as the CTC model, designed to reduce many aspects of scarcity (e.g., bed scarcity) and the unintended negative consequences (e.g., nosocomial infection) associated with inpatient care. In addition, the ability of CTC programs to treat large numbers of severely malnourished children as outpatients relies, to a large extent, on early detection and consequent early (low-dependency) treatment and resolution. For these reasons, the analysis of case-detection methods presented in this report will treat false positive misdiagnosis costs as being of secondary importance. It is important to note, however, that the requirement of moderate sensitivity at high specificity, as discussed above, will minimize the number of false positives.

Indicators of potential usefulness

Pelletier [18] identifies confusion between nutritional status and indicators of nutritional status as an additional source of controversy in selecting a case-detection method for malnutrition. The terms “nutritional status” and “anthropometric status” are, for example, often used interchangeably. Nutritional status refers to the internal state of an individual as it relates to the availability and utilization of nutrients at the cellular level. This state cannot be observed directly, so *observable indicators* are used instead. The range of indicators of nutritional status, none of which taken alone or in combination are capable of providing a full picture of an individual’s nutritional status, can be categorized as

» Biochemical: laboratory assays that measure specific aspects of a subject’s metabolism, such as tests to determine serum albumin levels:
» Clinical assessment: the presence of clinical signs suggestive of malnutrition, such as visible wasting and bipedal edema;
» Anthropometric: measurements of the physical dimensions of a subject used alone, in combination, or corrected for age.

Case definitions may use items from any or all of these categories (e.g., a case definition may use a single anthropometric indicator or use a diagnostic algorithm that combines biochemical tests, clinical assessment, and anthropometry).

Biochemical indicators require laboratory facilities, costly equipment, and highly qualified staff to perform and interpret tests, as well as equipment, facilities, and protocols for collecting, storing, and transporting specimens and for reporting results. These requirements make biochemical indicators unsuitable candidates for field-based case-detection methods. Case-detection methods using biochemical indicators will not, therefore, be considered further in this report.

A number of anthropometric indicators have been used in case definitions of severe malnutrition. This report considers weight-for-age (W/A), height-for-age (H/A), weight-for-height (W/H), mid-upper-arm circumference (MUAC), mid-upper-arm circumference-for-age (MUAC/A), and mid-upper-arm circumference-for-height (MUAC/H). In all cases the indicator is measured or derived from measured components (e.g., weight and height for W/H) and the value of the indicator is compared with a threshold value. Individuals for whom the indicator falls below the threshold value are classified as malnourished.

Considerations of how well a case definition may be said to represent an individual's nutritional status may not be the best criterion to judge the utility of a case definition in a programmatic context. Doing so may result in the selection of case definitions that are only weakly related to the aims of a program. The primary aim of most programs treating severe malnutrition is to prevent mortality. For such programs, therefore, the most useful case definition will be one that can identify individuals who are at high risk of dying if they remain untreated, but who would be likely to survive if treated in an appropriate nutritional support program. This realization has led a number of workers to argue that the utility of case definitions for malnutrition is defined more by their ability to reflect mortality risk than by their ability to reflect nutritional status [18–30].

This report will systematically review the relative utility of case definitions of severe malnutrition within the framework outlined in **table 1** and the preceding discussion.

Simplicity

Clinical assessment has proved successful with highly qualified clinical staff providing good reproducibility, validity (i.e., when compared with a range of biochemical indicators), and predictions of clinical course in surgical patients in a well-resourced setting [31]. Jelliffe and Jelliffe [16] caution that clinical assessment can only be performed by examiners who have been carefully and practically trained. Simoes et al. [32] reported good agreement between the clinical diagnosis of malnutrition made by trained nurses and by a reference pediatrician in primary-care settings in Ethiopia. Bern et al. [33] also reported good results with a single trained health worker in a district hospital in Kenya using visible severe wasting and/or bipedal edema as the case definition for severe malnutrition. This finding is, however, problematic, because anthropometric indicators (W/A and W/H) were used to validate the results, and the study subjects were weighed and measured and the anthropometric indicators were calculated at the time of the clinical assessment by the same health worker who performed the clinical assessment. Hamer et al. [34] reported poor results using the same case definition and validation criteria with trained registered and auxiliary nurses in a tertiary-level referral hospital in Gambia. In this study, the observers were initially blinded with regard to the anthropometric status of individual children.

Any indicator that includes an age component requires that age be ascertained accurately. Bairagi [35] reported that indicators that include an age component (i.e., H/A, W/A, and MUAC/A) are more sensitive to random errors in age than to random errors in anthropometry. Hamer et al. [34], working in a setting where accurate dates of birth were available, found that nurses had difficulty in accurately performing the arithmetic required to calculate age from date of birth and date of examination, although it should be noted that this was not covered in their training. Velzeboer et al. [36] reported that minimally trained community health volunteers in rural Guatemala had difficulties in performing date arithmetic.

Multicomponent indicators (i.e., W/A, H/A, W/H, MUAC/A, and MUAC/H) usually require finding values by looking them up in multidimensional tables or by plotting the values of the individual components on a "growth chart" for location with regard to a reference curve. This requires familiarity with a number of mathematical concepts (digit recognition, number formation, magnitude estimation, number order, number comparison, and graphical presentation of number), even if the required operations are to be performed mechanistically. Velzeboer et al. [36] tested the post-training ability of five minimally trained community health volunteers in rural Guatemala to calculate the W/H indicator. They reported that four of the five could not complete the test unsupervised because of problems with rounding decimal numbers (required for looking up values in tables) and that the one worker who completed the test unsupervised required over an hour to calculate 10 indicator values, of which 4 were incorrect. Hamer et al. [34] reported that registered and auxiliary nurses in a tertiary-level referral hospital in Gambia had difficulties in using growth charts immediately after training. It is unlikely, therefore, that these tasks could be performed by minimally trained community-based volunteers.

Sommer and Loewenstein [29] reported that MUAC/H, when measured with a device known as a QUAC stick, is a multicomponent indicator that does not require use of a table or reference to a growth chart. The QUAC (Quaker arm circumference) stick avoids the use of a table by having the MUAC thresholds defining malnutrition marked on a "height" stick. A child taller than the corresponding mark on the height stick for his or her measured MUAC is classified as malnourished. The impetus for the development of the QUAC stick was to improve the speed of measurement rather than to remove the need for supervision of staff during measurements. Davis [37] reported that under field conditions the method "was simple enough to be performed by unskilled Nigerians *under supervision*" (emphasis added). The utility, rapidity, and relative simplicity of the QUAC stick have also been reported by Loewenstein and Phillips [38] and Arnhold [39].

Alam et al. [19], in a comparison of W/A, H/A, W/H, MUAC, MUAC/A, and MUAC/H, reported that MUAC required only simple and inexpensive equipment and was faster and easier for minimally trained workers to perform in door-to-door screening than any of the other indicators tested. The fact that MUAC is a single linear measurement allows it to be used without the need for numbers, arithmetic, tables, or plotting of data on growth charts. Shakir and Morley [40] suggest the use of a color-banded cord to measure MUAC, with colors corresponding to classifications of malnutrition. Shakir [41] reported that a color-banded plastic strip simplified MUAC measurements further and provided immediate classifications in field situations when performed by minimally trained paramedical personnel in Iraq. This ability to make immediate classifications in the field by using a readily understandable “traffic light” system intuitively related to thinness may have a potential for raising awareness among community members of the prevalence of malnutrition, which is an essential first step in the process of mobilizing community action to counter the problem.

Acceptability

Velzeboer et al. [36], in a comparison of W/H and MUAC in Guatemala, reported that younger children tended to become upset and agitated during both weight and height measurements and that no such behavior was observed during the measurement of MUAC. Their characterization of these children as “traumatized” may be a little strong, as any trauma resulting from this situation is unlikely to have lasting consequences. The unpleasantness associated with weight and height measurement may, however, reduce the acceptability of indicators that use weight and/or height measurements to children, their carers, and community-based volunteers and have a negative impact upon the coverage of case-detection activities, particularly if carers of sick children refuse to have their children weighed and measured. Any tendency of younger children to become agitated during weight and height measurements may also have a negative impact on the precision and accuracy of measurement. There are no reports of difficulties in measuring height with the use of the QUAC stick.

Cost

Clinical assessment requires highly trained and relatively highly paid personnel if it is to be performed to an acceptable standard [16, 31, 34]. The opportunity costs associated with diverting clinic staff from direct patient care to community-based case-detection activities is a factor that should also be considered with regard to using clinical assessment for case detection in the community. Measurement of height and weight requires costly and delicate equipment that must be calibrated and maintained [29, 36, 37, 42]. The required equipment may not be available even at the level of the referral hospital [43]. The costs of providing and maintaining equipment may be acceptable in highly centralized programs with dedicated case-detection teams but are likely to prove unacceptable in programs relying on decentralized networks consisting of large numbers of community-based volunteers for case detection. Measurement of MUAC and MUAC/H by the QUAC stick can be performed with the use of low-cost and maintenance-free equipment [37, 40, 41]. To obtain weight and height measurements with precision and accuracy, it is generally considered that three persons are required: two to take the measurements and one to supervise, record the measurements, and calculate indicator values [44]. It may prove difficult to find a sufficient number of qualified community-based volunteers to undertake these measurements. The use of weight and/or height measurement will also have a considerable personnel, payroll, and logistics overhead if dedicated case-detection teams are employed.

Objectivity and quantitativeness

The subjective nature of clinical assessment may lead to acceptability problems, since carers may feel that nonclinical criteria (i.e., social, racial, or tribal discrimination) are being applied. Corruption is also an issue that must be considered with any subjective criterion. Clinical assessment is generally recognized as subjective, difficult to standardize, and difficult to express quantitatively [16, 34, 37]. Anthropometric indicators are both objective and quantitative, although there are problems of bias with indicators that include an age component when age cannot be ascertained accurately [34, 35].

Age independence

Age independence has two components. An indicator may be said to be independent of age if its value is not influenced by the age of the subject or if the predictive power (i.e., the power of predicting mortality) is independent of the age of the subject. One way of ensuring age independence is to adjust indicators to account for the age of the subject. This is done with H/A, W/A, and MUAC/A. The problem with this approach is that it is often difficult to ascertain age accurately [16, 34, 37], and indicators that include an age component are known to be more sensitive to random errors in age, which increase with increasing age, than to random errors in anthropometry [35]. In situations where the dates of birth or exact ages are unknown, this is likely to be a major problem. Because children grow fast, small errors in estimating age may lead to large errors in indicator values. In famine and in situations in which displacement and familial separation are common, fieldworkers are often required to estimate the age of children on the basis of little or no information. Estimates “by eye” are biased by assumptions about the relationship between height and age that are likely to be invalid in situations of nutritional stress. In these cases, indicator values will be subject to errors, probably systematic and upwards, that are products of random errors in estimating age and systematic errors in estimating age that may be influenced by growth failure [45]. MUAC and MUAC/H are known to be relatively independent of age, with reference medians increasing only slightly (i.e., by approximately 17 mm) between the ages of 1 and 5 years [16, 19, 30, 37, 42, 46, 47], but they are age-dependent in children below 1 year of age [47]. The relationship between MUAC and age is shown in **figure 1**. The predictive power of MUAC (i.e., the power of predicting mortality) is, however, independent of age even in children below 1 year of age [22, 30, 48–50]. Berkley et al. [50] reported consistently high case-fatality rates in hospitalized Kenyan children of all ages between 12 and 59 months with low MUAC values, which they define as ≤ 115 mm; this result suggests that unadjusted (i.e., by age) MUAC may be useful in clinical settings. W/H is also independent of age between the ages of 1 and 5 years [42, 51], but the predictive power (i.e., the power of predicting mortality) of W/H may change with age [26].

Precision and accuracy

The accurate ascertainment of age is problematic in many developing countries [16, 34, 37], which casts doubt on the accuracy of indicators that include an age component [35, 45]. It is often asserted that, in terms of precision and accuracy of measurement, MUAC compares unfavorably with W/H (e.g., Waterlow [51]). Evidence supporting such assertions is, however, elusive. Younger children tend to become agitated during weight and height measurement under field conditions [36]. This may have a negative impact on the precision and accuracy of height and weight measurements. Anthropometric indicators that include a height component assume that height cannot be lost.

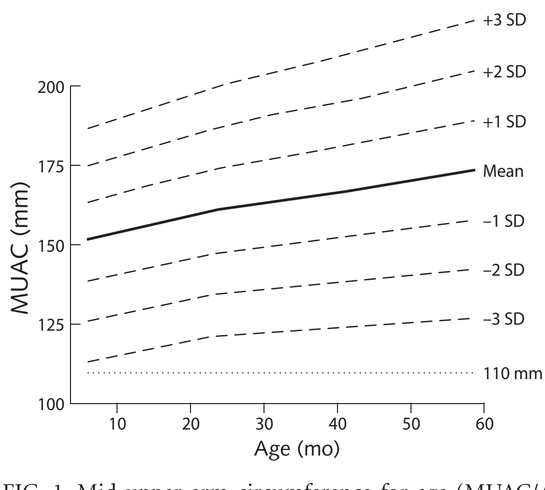

FIG. 1. Mid-upper-arm circumference-for-age (MUAC/A) growth reference curves for males and females aged between 6 and 59 months. MUAC/A growth reference curves presented in this figure are taken from de Onis et al. [47]

This assumption has not been tested in children, but it has been demonstrated to be invalid in adults in famine situations and in labor camps providing minimal "starvation" rations [45]. It should also be noted that weight may vary throughout the day, depending on factors such as hydration and the contents of the gastrointestinal tract, and that heavy parasitism with *Ascaris lumbricoides* may bias weight measurements upwards. Davis [37] reported that MUAC/H measured by a QUAC stick was both reproducible and accurate. This finding was confirmed by Sommer and Loewenstein [29]. Velzeboer et al. [36] tested the reliability (i.e., precision) of five minimally trained community health volunteers in rural Guatemala measuring W/H, H/A, W/A, MUAC, and MUAC/A. They reported that, under field conditions, intra-observer reliability was highest for W/A, followed by MUAC, MUAC/A, H/A, and W/H, and that inter-observer reliability was highest for W/A, followed by MUAC, MUAC/A, W/H, and H/A. Velzeboer et al. [36] also reported that under field conditions, minimally trained workers made fewer and smaller errors with MUAC than with W/A or W/H, even when they were not required to calculate indicator values by looking up values in tables or by plotting data on growth charts.

Feeney [9] reported that, with minimally trained community-based volunteers in a CTC program, the majority of errors were made in recording MUAC values (e.g., 104 mm recorded as 140 mm) rather than in deciding whether MUAC values fell above or below a threshold value. This study was undertaken in Ethiopia and required volunteers to work with a numbering system unfamiliar to them (using Roman rather than Amharic numerals). Recording errors did not have operational consequences, since referral for admission was determined by the subject's position with regard to a threshold value. A companion study found that when the volunteers were asked to classify children according to whether or not their MUAC fell below a fixed threshold of 110 mm, they made very few errors [9]. Feeney [9] and Spector [52] both identified pressure from carers to pull the MUAC strap tighter in order to facilitate admission as a source of a systematic downward bias in MUAC measurements made by community-based volunteers observed in a CTC program in Ethiopia. Such errors act to increase sensitivity at the cost of specificity.

Sensitivity, specificity, and predictive value

Loewenstein and Phillips [38] and Sommer and Loewenstein [29] reported that MUAC/H was strongly predictive of death at 1, 3, and 18 months after measurement. Kielmann and McCord [27] reported that W/A was predictive of death at 6 and 12 months after measurement in Indian children. Chen et al. [24] examined the associations between anthropometric indicators and subsequent mortality in Bangladeshi children. All indicators were negatively associated with mortality (i.e., the risk of death increased with decreasing values of the indicator). MUAC/A and W/A were the best predictors of death and W/H was the worst predictor. Trowbridge and Sommer [53], analyzing a subset of the data reported by Chen et al. [24], reported that MUAC alone performed better than MUAC/H and that MUAC adjusted for age (i.e., MUAC/A) was no more sensitive in relation to specificity than MUAC alone. Briend and Zimicki [22], using the same data as Sommer and Loewenstein [29] in a study to validate the use of MUAC as an indicator of risk of death within 1, 3, and 6 months of measurement in Bangladeshi children, reported that MUAC alone performed better in terms of both sensitivity and specificity than all other anthropometric indicators studied in the same and different populations. They confirm that correcting MUAC for age or height did little to improve sensitivity and specificity. This study demonstrates dramatic increases in sensitivity at high levels of specificity for shorter follow-up periods. In the context of case detection, short follow-up corresponds to frequent measurement, which is likely to be easier to achieve with simple, acceptable, and low-cost indicators measured by community-based volunteers than with less simple, less acceptable, and more expensive indicators measured by centralized screening teams [18]. Briend and Zimicki [22] examined the power of W/A, W/H, H/A, MUAC, and MUAC/A for predicting death in children hospitalized with diarrhea in a Dhaka hospital and reported that W/A, MUAC, and MUAC/A predicted death better than H/A and W/H. MUAC was the best univariate predictor of short-term mortality. This study also examined the possibility that combinations of indicators might have higher predictive power and found no combination of indicators that outperformed MUAC alone. Briend et al. [23] reported that MUAC, as an indicator of risk of death within 1 month of measurement in Bangladeshi children, was almost twice as sensitive as other anthropometric indicators at the same specificity and that only slight improvements in sensitivity could be achieved by using a diagnostic algorithm that used MUAC and selected clinical signs. Alam et al. [19], examining the use of MUAC, MUAC/A, MUAC/H, H/A, W/H, and H/A for predicting death 3 and 6 months after measurement in Bangladeshi children, reported that sensitivity at high levels of specificity was highest for MUAC and MUAC/A, intermediate for W/A, H/A, and MUAC/H, and lowest for W/H. Briend et al. [48] reported that MUAC without correction for age or height was superior in terms of sensitivity and specificity to W/A, H/A, and W/H in Senegalese children. Smedman et al. [28] reported that H/A, but not W/H, was a significant predictor of mortality in Bangladeshi children. Vella et al. [30] tested the predictive power of W/A, H/A, W/H, and MUAC in Ugandan children and found that in relation to specificity, MUAC was the most sensitive predictor of mortality within 12 months of measurement, followed by W/A, H/A, and W/H. In multivariate predictive models, MUAC was found to increase the predictive power of other indicators, whereas other indicators did not improve the predictive power of MUAC. Berkley et al. [49] reported that MUAC and W/H had similar predictive power with regard to mortality in a large inpatient cohort of Kenyan children. In summary, the most consistently reported observation is that W/H is the least effective predictor of mortality and that, at high specificities, MUAC is superior to H/A and W/A.

Marasmus and kwashiorkor

A problem with relying on a single anthropometric indicator for malnutrition is that the predominant form of severe malnutrition is marasmus in some contexts and kwashiorkor in others [16]. This problem is usually addressed by using an anthropometric indicator to define marasmus and the presence or absence of bipedal edema to define kwashiorkor [51]. Kahigwa et al. [54] reported substantial agreement between two clinical officers in a Tanzanian hospital for identification of edema. Hamer et al. [34] reported that trained registered and auxiliary nurses in a tertiary-level referral hospital in Gambia performed poorly at identifying bipedal edema, and it was observed that the nurses spent insufficient time depressing tissues. Simoes et al. [32] reported good agreement between the clinical diagnosis of malnutrition made by trained nurses and by a reference pediatrician in primary-care settings in Ethiopia. This suggests that, as with all clinical assessment, careful and practical training of workers is required to achieve reasonable levels of sensitivity and specificity for detecting cases of kwashiorkor.

W/H-based indicators used alone (i.e., without examination for bipedal edema) are poor at detecting cases of kwashiorkor, because the weight of retained fluid tends to mask what would otherwise be low W/H values. Sandiford and Paulin [55] reported that MUAC used alone was more sensitive and more specific than either W/H and W/A used alone as a test for bipedal edema in Malawi. Berkley et al. [49] reported that MUAC used alone performed better than W/H used alone at identifying children with bipedal edema and skin and hair changes associated with kwashiorkor in Kenya. Currently available data suggest that the use of MUAC may, to some extent, compensate for the potentially poor performance of minimally trained community-based volunteers in identifying bipedal edema by clinical examination.

The use of anthropometry in young children

Anthropometric measurements are difficult to perform on young children. Children under 6 months of age weigh only a few kilograms. To obtain sufficiently accurate measurements of weight, children aged less than 6 months should be weighed on specialist pediatric scales that are graduated in units of 10 g rather than on conventional hanging scales that are graduated in units of 100 g. This requires the provision and maintenance of suitable scales. The length of children less than 6 months old can be measured with conventional height boards, but very small infants are difficult to handle and great care needs to be exercised when measuring them. For these reasons, admission of younger children to therapeutic feeding programs tends to be based on subjective criteria, such as visible severe wasting and assessments of risk factors. The use of MUAC in this context is also problematic, since, in contrast to older children, there are no data suggesting an association between MUAC and mortality that is independent of age in this age group. Moreover, internationally recognized reference curves remain unavailable for this age group [47].

The use of anthropometry in adolescents

The use of anthropometry in adolescents is subject to similar problems as in young children. Weight measurement in adolescents requires physician scales. Height measurement in adolescents requires height boards capable of measuring heights of 2 m or above. This requires the provision and maintenance of suitable scales and height boards. The interpretation of anthropometric measures in adolescents is complicated by changes in body shape, body composition, and musculature that occur during puberty. The use of MUAC without correction for age in this age group is also problematic due to changes in musculature during puberty and because, in contrast to younger age groups, there are no data suggesting an association between MUAC and mortality that is independent of age in this age group. Adjusting MUAC for age is likely to be needed in this age group.

Summary

Table 2 summarizes the data presented above according to whether specific indicators exhibit the key properties outlined in the conceptual and methodologic framework. Within this framework, MUAC or MUAC/H measured with the QUAC stick plus the presence of bipedal edema are the indicators most suited to screening and case detection for malnutrition in the community. MUAC/H appears to offer no significant advantage over MUAC alone, which is the simpler and cheaper measure. There also remains some doubt as to whether the QUAC stick can be used by minimally trained community-based volunteers without supervision. It is important to note that W/H, which is the commonest indicator used for screening and case detection of malnutrition in the community, is, when reviewed within the conceptual and methodologic framework used in this report, one of the least useful indicators in this context.

The fact that MUAC is simple, objective, quantitative, precise, and accurate means that a referral by a community-based volunteer can be treated as an admission entitlement, with all referrals automatically admitted upon presentation of a valid referral slip. Referral slips can be numbered in such a way as to identify the source of referral and prevent fraud. Suitable books of slips are already available at low cost and are sold as "cloakroom

TABLE 2. Capability of common indicators with regard to key properties of case-detection methods for screening and case detection of malnutrition in the community

	Indicator						
Property	Clinical	W/A	H/A	W/H	MUAC	MUAC/A	MUAC/H
Simplicity	No	No	No	No	Yes	No	Yes (by QUAC stick only)
Acceptability	No	No	No	No	Yes	Yes	Yes (by QUAC stick only)
Cost	No	No	No	No	Yes	Yes	Yes (by QUAC stick only)
Objectivity	No	No	No	Yes	Yes	No	Yes
Quantitativeness	No	Yes	Yes	Yes	Yes	Yes	Yes
Independence of age	Yes	No	No	No	Yes	No	Yes
Precision (reliability)	No	Yes	No	No	Yes	Yes	Yes (by QUAC stick only)
Accuracy	No	No	No	No	Yes	No	Yes
Sensitivity	NA	Yes	No	No	Yes	Yes	Yes
Specificity	NA	Yes	No	No	Yes	Yes	Yes
Predictive value	NA	Yes	No	No	Yes	Yes	Yes

W/A, weight-for-age; H/A, height-for-age; W/H, weight-for-height; MUAC, mid-upper-arm circumference; MUAC/A, mid-upper-arm circumference-for-age; MUAC/H, mid-upper-arm circumference-for-height; QUAC, Quaker arm circumference tickets" or "raffle tickets (see **fig. 2**). Remeasurement of MUAC at admission will allow such a referral and admission system to be monitored in order to identify problems with particular volunteers. Since each referral and admission has a unique number that can identify the source of referral, and case finders have a defined catchment area, it would be relatively easy to monitor did-not-attend (DNA) rates through a routine admissions-monitoring system.

Selecting an appropriate indicator threshold

Using an anthropometric indicator such as MUAC in case definitions of malnutrition requires that the indicator be measured and the value of the indicator compared with a threshold value. Individuals for whom the indicator falls below the threshold value are classified as malnourished. With banded MUAC straps such as those proposed by Shakir and Morley [40] and Shakir [41], the threshold can be color-coded on the strap, providing a simple-to-use, instantaneous, and unambiguous indicator as to whether a child falls above or below the case-defining threshold.

The factors that influence the choice of threshold value are the sensitivities, specificities, and predictive values for mortality associated with threshold values. **Figure 3** shows the relationship between MUAC and mortality, expressed in deaths per 1,000 child-years, as reported in separate studies by Briend and Zimicki [22], Briend et al. [23], Alam et al. [19], Pelletier et al. [56], and Vella et al. [30]. Mortality increases exponentially with declining MUAC, with small increases in mortality at intermediate MUAC values (i.e., between 110 and 130 mm) and large increases in mortality at MUAC values below 110 mm. There is little between-study variation in the observed relationships, despite the fact that these studies were undertaken by different teams in different locations at different times, with varying lengths of follow-up and inconsistent censoring of accidental deaths. The available data on the relationship between MUAC and mortality suggest that there is little justification in setting the case-defining threshold below about 110 mm. As shown in **figure 1**, this threshold is equal to or more extreme than 3 z-scores below the mean of the sex-combined MUAC/A reference distribution for children aged 7 months or older and equal to or more extreme than 4 z-scores below the mean of the sex-combined MUAC/A reference distribution for children aged 39 months or older [47].

A proposed case definition

Currently available data suggest that the case definition

MUAC < 110 mm OR the presence of bipedal edema,

with MUAC measured with the use of color-banded straps, is suitable for use by minimally trained community-based volunteers with limited schooling and low levels of numeracy and literacy.

It should be noted that this proposed case definition applies only to children aged between 6 months

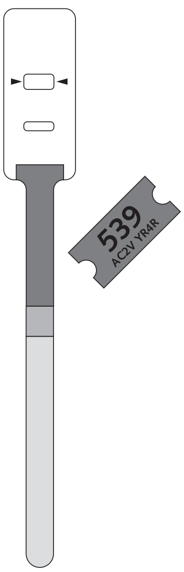

FIG. 2. Banded mid-upper-arm circumference (MUAC) strap and cloakroom/raffle ticket referral slip

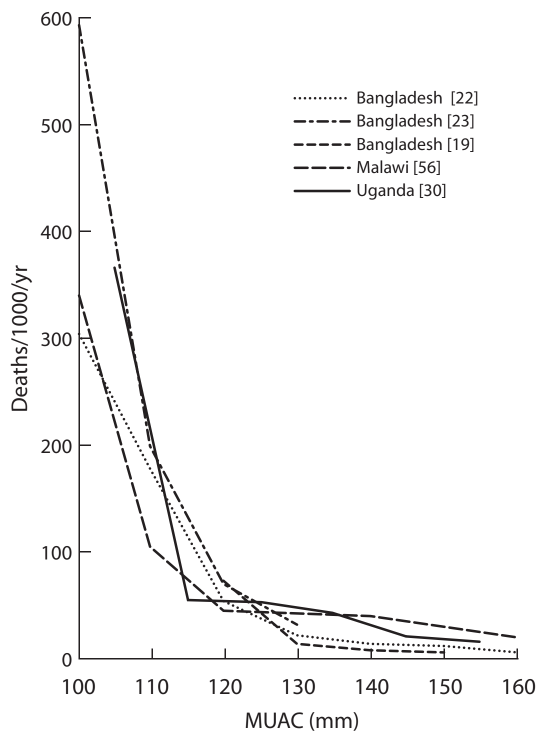

FIG. 3. Observed relationship between mid-upper-arm circumference (MUAC) and child mortality in five studies: Briend and Zimicki [22], Briend et al. [23], Alam et al. [19], Pelletier et al. [56], and Vella et al. [30]

and 5 years. Height may be used as a proxy for age. In this case, the proposed case definition applies only to children between 65 and 110 cm in height, with eligibility ascertained by a simple marked stick. These height thresholds are conventional and may not be appropriate in settings where infantile stunting is common. In such settings, local H/A data could be used to decide suitable height thresholds.

Triage, response, and appropriate resource utilization

The primary aim of most programs treating severe malnutrition is to prevent mortality. For such programs, therefore, the most useful case definition will be one that can identify individuals who are at high risk of dying if they remain untreated but would be likely to survive if treated in an appropriate nutritional support program. Currently available data indicate that MUAC is one of the best predictors of mortality, but children selected for treatment because they have extremely low values of MUAC may die even when treated. Admitting such children would then be an inappropriate use of resources. The use of a MUAC case definition should, therefore, be examined with regard to clinical triage. The triage categories and outcomes for programs treating malnutrition are shown in **table 3**.

The intensity of intervention that is required for children with extremely low values of MUAC is also of interest. If children with extremely low values of MUAC do well when treated with low-intensity interventions, such as being admitted to a supplementary feeding program (SFP), then treating them with a comparatively high-intensity intervention, such as therapeutic feeding in an OTP, would be an inappropriate use of resources. This question is of particular interest in smaller children, usually defined as those under 12 months of age or of height ≤ 75 cm (i.e., the approximate H/A reference median for 12-month-old children), where the use of case definitions based on unadjusted (i.e., for age or height) MUAC values is the cause of some controversy.

The two questions of interest for CTC implementation are the following:

» Do smaller children with extremely low values of MUAC do well in OTP?
» Do smaller children with extremely low values of MUAC do well in SFP?

TABLE 3. Triage categories for programs treating malnutrition

Triage category	Response to intervention	Triage outcome
Not malnourished	Intervention not indicated	Do not admit
Malnourished (treatable)	Will benefit from intervention	Admit
Malnourished (untreatable)	Will not benefit from intervention	Do not admit

A natural experiment in a CTC program in Northern Ethiopia in 2003 provides answers to these questions for smaller children without bipedal edema and with a W/H greater than 70% of the median of the reference population. When this program started in February 2003, children with the case definition MUAC < 110 mm AND (age > 12 months OR height > 75 cm) AND W/H > 70% were admitted to the OTP. In March 2003, the case definition was changed, on the strong advice of an acknowledged international expert on malnutrition, to MUAC < 110 mm AND height > 75 cm AND W/H > 70%. The effect of this change was to exclude, among children with MUAC below 110 mm, the smaller ones (i.e., those whose height was ≤ 75 cm) from admission to the OTP. This change in case definitions created a natural experiment with two comparable groups of children with MUAC below 110 mm, with height ≤ 75 cm, with W/H greater than 70% of the reference median, and without bipedal edema being admitted initially to OTP and then to SFP. This was noted during a program review in November 2003 and allowed a comparison of the responses of smaller children with extremely low values of MUAC admitted to OTP and SFP. Summary data from the natural experiment are presented in **table 4**.

There is some doubt regarding the accuracy of age reporting in the OTP arm of the natural experiment. Examination of the individual records together with the similarity in the distributions of heights between the two groups suggests preferential reporting of age as 13 months in the OTP arm. This may have been due to deliberate misreporting of age by carers or deliberate misrecording of age by program staff in order to facilitate admission of younger children into the more

TABLE 4. Summary of data arising from a natural experiment allowing comparison of response to treatment of children with MUAC < 110 mm, height ≤ 75 cm, W/H > 70% of the reference median, and without edema in OTP and SFP

	Experimental arm	
Variable	OTP	SFP
No. of subjects	42	56
No. of survivors	40	46
No. of deaths	0	8
No. lost to follow-up or defaulted	2	2
Age range (median)	12–36 mo (16 mo)	6–36 mo (14 mo)
Height range (median)	62–72 cm (66 cm)	54–75 cm (67 cm)
MUAC range (median)	82–109 mm (104 mm)	85–109 cm (102 mm)
Sex ratio	54% male	57% male

MUAC, mid-upper-arm circumference; W/H, weight-for-height; OTP, outpatient treatment program; SFP, supplementary feeding program intensive OTP program. It is likely, therefore, that the distributions of ages are similar in both arms of the natural experiment.

Table 5 shows a crude analysis of the survival data in the two arms of the natural experiment. The effect observed in this crude analysis remains statistically significant after adjustment for age at admission split into less than 13 months of age and 13 months of age or older (Mantel-Haenszel $\chi^2 = 3.86$, df = 1, $p = .0494$). This analysis is compromised by probable inaccurate reporting and/or recording of age. The effect observed in the crude analysis remains statistically significant after adjustment for height (as a proxy for age) at admission split into above or below the overall median height at admission of 66.15 cm (Mantel-Haenszel $\chi^2 = 4.89$, df = 1, $p = .0269$).

Figure 4 shows the results of an analysis of weight gains in grams per kilogram per day observed in the two arms of the natural experiment. Smaller children with MUAC less than 110 mm responded well (in terms of both survival and weight gain) to the high-intensity intervention (OTP) but did not respond well to the low-intensity intervention (SFP). Treating such children with a high-intensity intervention such as therapeutic feeding in an OTP is likely, therefore, to be an appropriate use of resources. The findings of this natural experiment suggest that smaller children (i.e., those aged below 12 months or whose height is ≤ 75 cm) with MUAC < 110 mm should be admitted to programs treating severe malnutrition.

It should be noted that the two arms of the natural experiment were sequential rather than concurrent. It is possible, therefore, that the observed differences were due, in some part, to seasonal factors such as changes in the incidence of malaria. The protocol for the OTP included weekly examination by a clinical officer as well as systematic treatment with antibiotics and malaria prophylaxis at the start of the treatment episode. None of these services were provided by the SFP. If children during the later (SFP) arm of the study had been admitted to OTP, they would, therefore, have been considerably more likely to receive timely and appropriate treatment and prophylaxis. The OTP arm ran during the period of high malaria incidence following the short (*Belg*) rains. The SFP arm ran for 7 months, with 2 months during the period of high malaria incidence at the end of and following the long (*Meher*) rains. It is likely, therefore, that the differences observed in the natural experiment were due, in large part, to differences in program intensity rather than to seasonal factors.

TABLE 5. Crude analysis of survival data from a natural experiment allowing comparison of response to treatment of children with MUAC < 110 mm, height ≤ 75 cm, W/H > 70% of the reference median, and without edema in OTP and SFP

	Outcome		
Exposure	Died	Survived	Total
SFP	8	46	54
OTP	0	40	40
Total	8	86	94

MUAC, mid-upper-arm circumference; W/H, weight-for-height; OTP, outpatient treatment program; SFP, supplementary feeding program

Fisher-Irwin exact test: $p = .0094$ (one-sided); $p = .0191$ (two-sided)
Risk difference = 14.81%; 95% confidence interval, 3.15% to 26.47%
z-test: $z = 2.17$
$p = .0149$ (one-sided); $p = .0299$ (two-sided)

Implications of changing to MUAC-based case-selection methods

The most commonly used case definition for therapeutic feeding programs is W/H < 70% of reference median OR the presence of bipedal edema. Changing this to MUAC < 110 mm OR the presence of bipedal edema may have significant implications for program size, particularly in contexts where marasmus is the predominant form of severe malnutrition. Anecdotal evidence from Ethiopian CTC programs suggests that use of the MUAC-based case definition is likely to result in larger programs than use of the W/H-based case

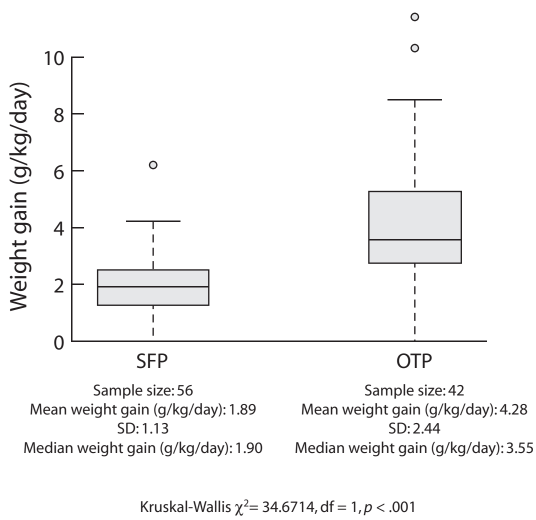

FIG. 4. Observed weight gains (g/kg/day) from a natural experiment allowing comparison of response to treatment of children with mid-upper-arm circumference (MUAC) < 110 mm, height ≤ 75 cm, W/H > 70% of the reference median, and without edema in outpatient treatment programs (OTP) and supplementary feeding programs (SFP).

W/H weight-for-height; MUAC, mid-upper-arm circumference; OTP, outpatient therapeutic program; SFP, supplementary feeding program

The central horizontal line in the boxes represents the median; the ends of the central boxes represent the upper and lower quartiles; the "whiskers" extend to 1.5 times the interquartile range; and the plotted points represent outliers.

definition. This was tested by a simple computer-based simulation. More than 200 datasets from nutritional anthropometry surveys that collected data on sex, weight, height, MUAC, and edema were obtained from international nongovernmental organizations. These datasets were restructured to ensure compatible coding between them and combined into a single large dataset representing more than 210,000 children between 65 and 110 cm in height. The prevalence of malnutrition in the combined dataset according to standard W/H-based case definitions is summarized in **table 6**.

The following case definitions were applied to all children in the combined dataset:

The W/H-based case definition was (MUAC < 125 mm AND W/H < 70% of the reference median) OR edema. The MUAC-based case definition was MUAC < 110 mm OR edema. The W/H-based case definition includes a MUAC measurement in order to simulate a two-stage screening procedure with a reasonably sensitive MUAC screen as the first screening stage. **Figure 5** shows in graphical form the result of applying these case definitions to the combined dataset. The MUAC-based case definition resulted in a larger program than the W/H-based case definition:

» Number of malnourished children identified by the MUAC-based case definition: 5,484;
» Number of malnourished children identified by the W/H-based case definition: 3,678.

The overall need in the combined dataset was estimated as the number of children identified as severely malnourished by either case definition [57]. The MUAC-based and W/H-based case definitions selected many of the same children. When the MUAC-based case definition is used, the number of excluded low-W/H children is small relative to estimated overall need:

» Estimated overall need (i.e., number of children identified as malnourished by either the MUAC-based or the W/H-based case definition): 5,867;
» Number of malnourished children excluded by the MUAC-based case definition: 383 (6.53%).

When the W/H-based case definition is used, however, the number of excluded low-MUAC children is large relative to the estimated overall need:

» Estimated overall need (i.e., number of children identified as malnourished by either the MUAC-based or the W/H-based case definition): 5,867;
» Number of malnourished children excluded by the W/H-based case definition: 2,189 (37.31%).

Figure 6 shows the age profiles of the children *excluded* by the two case definitions. The age profile of the excluded low-W/H children differs from the age profile of the excluded low-MUAC children. The children excluded by the W/H-based case definition tend to be younger and, hence, at higher risk of mortality than those excluded by the MUAC-based case definition.

These results assume programs with 100% coverage of case-finding activities and 100% uptake of services. Such assumptions are unrealistic, since no case-finding method is likely to achieve 100% coverage of case-finding activities, and no program is likely to achieve 100% uptake. Case-finding activities using a MUAC-based case definition are likely to have a higher coverage (as a result of simplicity, acceptability, low cost, and effective use of community-based volunteers and program staff) than case-finding activities using a W/H-based case definition. Programs using a MUAC-based case definition are likely to have a higher uptake (as a result of minimizing the problems of rejected referrals, crowding, and long waiting times) than programs using a W/H-based case definition. The results presented in **figure 5** are, therefore, subject to considerable bias. The relative difference in the sizes of the two programs is likely to be larger, the proportion of children excluded by the W/H-based case definition is likely to be larger, and the proportion of children excluded by the MUAC-based case definition is likely to be smaller than the figures presented in **figure 5** suggest.

Adopting a MUAC-based case-detection method will require changes to the way epidemiologic and needs-

TABLE 6. Prevalence of malnutrition in the combined dataset according to case definitions based on W/H z-score and the presence or absence of bipedal edema

Malnutrition category	Case definition	Prevalence
Global	W/H z-score < –2.00 and/or bipedal edema	11.70%
Moderate	W/H z-score < –2.00 without bipedal edema	9.10%
Severe (marasmus)	W/H z-score < –3.00 without bipedal edema	1.30%
Severe (kwashiorkor)	Bipedal edema	1.30%

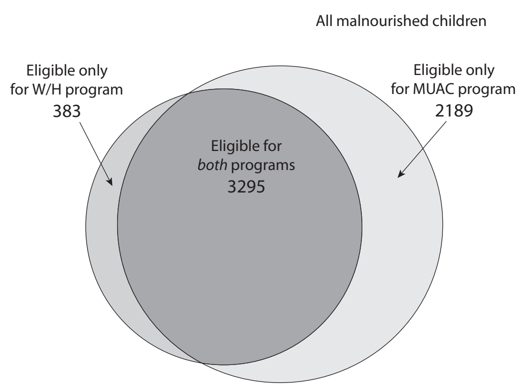

FIG. 5. Number of children selected from the combined dataset using two different case definitions. Weight-for-height (W/H) program: (MUAC < 125 mm AND W/H < 70% of the reference median) OR edema. Mid-upper-arm circumference (MUAC) program: MUAC < 110 mm OR edema. Overall need (labeled as "All malnourished children") is the number of children identified by either case definition.

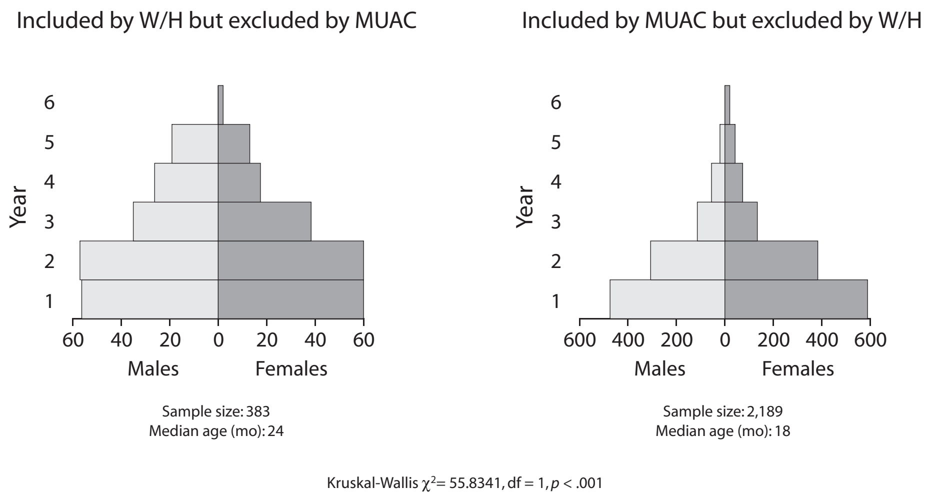

FIG. 6. Age profiles of children excluded by two different case definitions. W/H, weight-for-height; MUAC, mid-upper-arm circumference

assessment surveys are carried out. At present, these surveys estimate prevalence and need using slightly different variants of the W/H indicator. As need becomes defined by MUAC rather than by W/H. these surveys will need to collect MUAC in addition to weight and height for the purpose of needs estimation.

Monitoring and discharge criteria

Data from the natural experiment in an Ethiopian CTC program demonstrate that MUAC does respond to treatment (**figure 7**), but there are no good reasons to assume that an indicator that is suited to case detection will also be well suited to monitoring the progress of patients in a program or for deciding whether or not a patient may be discharged from a program [17]. At present there are no compelling data to suggest a move away from a weight-based indicator toward a MUAC-based indicator for monitoring and discharge. It should be noted, however, that height boards are often unavailable in primary health-care centers in developing countries. This means that using W/H for monitoring and discharge is problematic. Retaining W/H for monitoring and discharge also raises the problem that some children will be admitted on the basis of MUAC who are already above the W/H discharge criterion. Current practice in CTC programs for such cases is to monitor weight and to

» Discharge a patient as *cured* after a minimum of 2 months in OTP if MUAC > 110 mm, edema has been absent for a minimum of 2 weeks, sustained weight gain has occurred, and the patient is “clinically good”;
» Discharge a patient as a *nonresponder* after a minimum of 4 months in OTP if weight is stable and all available treatment options (e.g., home visits, inpatient stabilization, hospitalization, antiretroviral treatment (ART) programs, and tuberculosis treatment programs) have been pursued.

These monitoring and discharge criteria may be applied to all cases. The advantage of this approach is that it requires that only weight be monitored, and suitable scales are usually available in primary health-care centers in developing countries that have growth-monitoring programs. Monitoring weight alone does

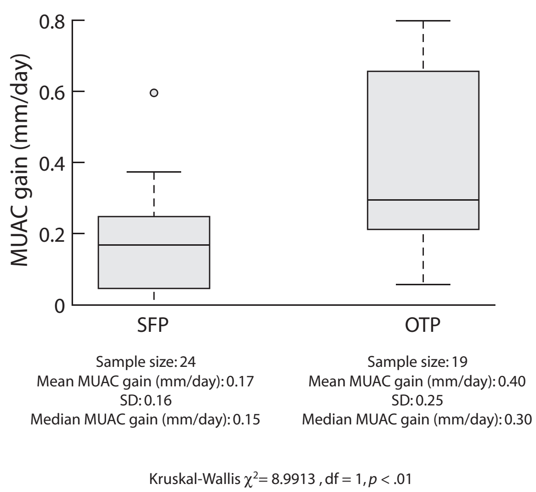

FIG. 7. Observed mid-upper-arm circumference (MUAC) gains (mm/day) from a natural experiment allowing comparison of response to treatment of children with MUAC < 110 mm, height ≤ 75 cm, weight-for-height (W/H) > 70% of the reference median, and without edema in outpatient treatment programs (OTP) and supplementary feeding programs (SFP).

The central horizontal line in the boxes represents the median; the ends of the central boxes represent the upper and lower quartiles; the “whiskers” extend to 1.5 times the interquartile range; and the plotted points represent outliers.

not differ greatly from monitoring W/H, because height changes little during recovery, and changes in W/H are due mainly to changes in weight rather than height; and because when W/H is monitored, a single height measurement, usually taken at admission, is often used throughout the treatment episode.

An alternative approach that also requires that only weight be monitored would be to use percentage weight gain:

$$\frac{\text{Current weight} - \text{Weight at admission}}{\text{Weight at admission} \times 100}$$

as a discharge criterion. With this approach, patients would be discharged once their percentage weight gain exceeded a cutoff value based on their weight at admission (or weight at loss of edema for patients presenting with marasmic kwashiorkor). Preliminary analysis of data from CTC programs in Malawi and Ethiopia suggests that a cutoff of 15% would result in approximately 50% of discharges meeting or exceeding 80% of the W/H reference median, and that a cutoff of 18% would result in approximately 50% of discharges meeting or exceeding 85% of the W/H reference median. Percentage weight gain could be combined with a MUAC cutoff. For example:

> Discharge as cured if MUAC ≥ 115 mm AND percentage weight gain ≥ 15%

The calculation of percentage weight gain could be simplified by the use of a lookup table. For example, **table 7** shows discharge weights for admission weights based on a 15% weight gain.

There are aspects of CTC programs (e.g., the concentration on maximizing program coverage and community-based delivery of services) that are more typical of “public health” or “mass treatment” inter-

TABLE 7. Example of a look-up table for calculating percentage change (15% in this table) in weight (kg)

Weight at		Weight at		Weight at		Weight at		Weight at		Weight at		Weight at	
Admission	Discharge	Admission	Discharge	Admission	Discharge	Admission	Discharge	Admission	Discharge	Admission	Discharge	Admission	Discharge
4.0	4.6	7.0	8.1	10.0	11.5	13.0	15.0	16.0	18.4	19.0	21.9	22.0	25.3
4.1	4.7	7.1	8.2	10.1	11.6	13.1	15.1	16.1	18.5	19.1	22.0	22.1	25.4
4.2	4.8	7.2	8.3	10.2	11.7	13.2	15.2	16.2	18.6	19.2	22.1	22.2	25.5
4.3	4.9	7.3	8.4	10.3	11.8	13.3	15.3	16.3	18.7	19.3	22.2	22.3	25.6
4.4	5.1	7.4	8.5	10.4	12.0	13.4	15.4	16.4	18.9	19.4	22.3	22.4	25.8
4.5	5.2	7.5	8.6	10.5	12.1	13.5	15.5	16.5	19.0	19.5	22.4	22.5	25.9
4.6	5.3	7.6	8.7	10.6	12.2	13.6	15.6	16.6	19.1	19.6	22.5	22.6	26.0
4.7	5.4	7.7	8.9	10.7	12.3	13.7	15.8	16.7	19.2	19.7	22.7	22.7	26.1
4.8	5.5	7.8	9.0	10.8	12.4	13.8	15.9	16.8	19.3	19.8	22.8	22.8	26.2
4.9	5.6	7.9	9.1	10.9	12.5	13.9	16.0	16.9	19.4	19.9	22.9	22.9	26.3
5.0	5.8	8.0	9.2	11.0	12.7	14.0	16.1	17.0	19.6	20.0	23.0	23.0	26.5
5.1	5.9	8.1	9.3	11.1	12.8	14.1	16.2	17.1	19.7	20.1	23.1	23.1	26.6
5.2	6.0	8.2	9.4	11.2	12.9	14.2	16.3	17.2	19.8	20.2	23.2	23.2	26.7
5.3	6.1	8.3	9.5	11.3	13.0	14.3	16.4	17.3	19.9	20.3	23.3	23.3	26.8
5.4	6.2	8.4	9.7	11.4	13.1	14.4	16.6	17.4	20.0	20.4	23.5	23.4	26.9
5.5	6.3	8.5	9.8	11.5	13.2	14.5	16.7	17.5	20.1	20.5	23.6	23.5	27.0
5.6	6.4	8.6	9.9	11.6	13.3	14.6	16.8	17.6	20.2	20.6	23.7	23.6	27.1
5.7	6.6	8.7	10.0	11.7	13.5	14.7	16.9	17.7	20.4	20.7	23.8	23.7	27.3
5.8	6.7	8.8	10.1	11.8	13.6	14.8	17.0	17.8	20.5	20.8	23.9	23.8	27.4
5.9	6.8	8.9	10.2	11.9	13.7	14.9	17.1	17.9	20.6	20.9	24.0	23.9	27.5
6.0	6.9	9.0	10.4	12.0	13.8	15.0	17.3	18.0	20.7	21.0	24.2	24.0	27.6
6.1	7.0	9.1	10.5	12.1	13.9	15.1	17.4	18.1	20.8	21.1	24.3	24.1	27.7
6.2	7.1	9.2	10.6	12.2	14.0	15.2	17.5	18.2	20.9	21.2	24.4	24.2	27.8
6.3	7.2	9.3	10.7	12.3	14.1	15.3	17.6	18.3	21.0	21.3	24.5	24.3	27.9
6.4	7.4	9.4	10.8	12.4	14.3	15.4	17.7	18.4	21.2	21.4	24.6	24.4	28.1
6.5	7.5	9.5	10.9	12.5	14.4	15.5	17.8	18.5	21.3	21.5	24.7	24.5	28.2
6.6	7.6	9.6	11.0	12.6	14.5	15.6	17.9	18.6	21.4	21.6	24.8	24.6	28.3
6.7	7.7	9.7	11.2	12.7	14.6	15.7	18.1	18.7	21.5	21.7	25.0	24.7	28.4
6.8	7.8	9.8	11.3	12.8	14.7	15.8	18.2	18.8	21.6	21.8	25.1	24.8	28.5
6.9	7.9	9.9	11.4	12.9	14.8	15.9	18.3	18.9	21.7	21.9	25.2	24.9	28.6

ventions than traditional center-based models of service delivery. In such interventions, adherence to stringent technical standards, service delivery, and the achievement of high coverage takes precedence over individual responses to the delivered intervention. From this perspective, it may be reasonable to adopt a fixed length of treatment episode for CTC programs. This approach does not differ much from current practice in programs using W/H or edema for admission. In such programs, patients admitted with edema but with a W/H percentage of median above 80% are, typically, retained in the program for a fixed period after loss of edema. Preliminary analysis of data from CTC programs in Malawi and Ethiopia suggests that an episode length of 60 days would result in approximately 50% of discharged patients achieving a weight gain of at least 15% at discharge.

As data from CTC programs become available, it will be possible to refine discharge criteria.

Conclusions

Within the framework of analysis adopted for this report, subjective clinical assessment (i.e., visible severe wasting) performs worse than any anthropometry-based method. W/H-based case-detection methods perform worse (i.e., in terms of age independence, precision, accuracy, sensitivity, and specificity) than any alternative anthropometry-based method and are neither simple, cheap, nor acceptable. Currently available evidence indicates that MUAC is the best (i.e., in terms of age independence, precision, accuracy, sensitivity, and specificity) case-detection method for severe malnutrition and that it is also simple, cheap, and acceptable. It is recommended, therefore, that programs treating severe malnutrition move toward MUAC-based case-detection, referral, and admission criteria. There is no compelling evidence supporting a move away from using weight in combination with clinical criteria (e.g., loss of edema) for monitoring and discharge.

References

1. Van Damme W. Medical assistance to self-settled refugees. Guinea 1990–1996. Antwerp, Belgium: ITG Press, 1998.
2. Van Damme W, Boelaert M. Medical assistance to self-settled refugees in Guinea. Lancet 2002; 359:260–1.
3. Collins S. Changing the way we address severe malnutrition during famine. Lancet 2001; 358(9280):498–501.
4. World Health Organization. Management of severe malnutrition: a manual for physicians and other senior health workers. Geneva: WHO, 1999. Available at: http://www.who.int/nutrition/publications/en/manage_severe_malnutrition_eng.pdf. Accessed 28 May 2006.
5. Sphere Project. The Sphere humanitarian charter and minimum standards in disaster response, 2nd ed (summer 2003 draft). Geneva: Sphere Project, 2003.
6. Collins S. Community-based therapeutic care: a new paradigm for selective feeding in nutritional crises. Network Paper No. 48. London: Humanitarian Practice Network, 2004.
7. Khar T. Report on focus group discussions: investigation into beneficiary opinions of the CTC programme at distribution sites (internal report). Dublin: Valid International/ Worldwide, 2003.
8. Golden K, Khara T. CTC in Ethiopia: working from CTC principles. In: Collins S, Khara T, eds. Community-based therapeutic care (CTC). Special Supplement No. 2. Oxford, UK: Emergency Nutrition Network, 2004.
9. Feeney B. Investigation into community volunteers using an admission criteria of middle upper arm circumference of below 110 mm and length equal to or greater than 65 cm for children to a community programme for severely malnourished children in Ethiopia. MSc thesis, London School of Hygiene and Tropical Medicine, London, 2004.
10. Concern Worldwide, Médecins sans Frontières Switzerland, Médecins sans Frontières USA. Community-based therapeutic care (CTC) in Ethiopia. Proceedings of a workshop held in Addis Ababa, Ethiopia, June 22–23, 2004.
11. Khara T, Collins S. Community based therapeutic care. Emergency Nutrition Network Supplement Series No. 2. Oxford, UK: Emergency Nutrition Network, 2004.
12. Lee J. Improving the community-based therapeutic care programme coverage, Dowa District, Malawi (internal report). Oxford, UK: Valid International, 2002.
13. Bandawe CR, Kabwazi N. The cultural and social factors affecting the implementation and success of the Community-Based Therapeutic Care (CTC) approach to treating severe acute malnutrition in Dowa, Malawi (internal report). Oxford, UK: Valid International, 2003.
14. Sackett DL, Holland WW. Controversy in the detection of disease. Lancet 1975; 2(7930):357–9.
15. Beaton GH, Bengoa JM. Practical population indicators of health and nutrition. In: Beaton G, Bengoa JM, eds. Nutrition and preventive medicine. Monograph Series 62. Geneva: World Health Organization, 1976:500–19.
16. Jelliffe EFP, Jelliffe DB. The arm circumference as a public health index of protein-calorie malnutrition of early childhood. J Trop Pediatr 1969; 15:179–92.
17. Habicht JP. Some characteristics of indicators of nutritional status for use in screening and surveillance. Am J Clin Nutr 1980; 33:531–5.
18. Pelletier DL. The relationship between child anthropometry and mortality in developing countries: implications for policy, programs and future research. J Nutr 1994; 124(10 suppl):2047S–81S.
19. Alam N, Wojtyniak B, Rahaman MM. Anthropometric indicators and risk of death. Am J Clin Nutr 1989; 49: 884–8.
20. Bairagi R, Chowdhury MK, Kim YJ, Curlin GT. Alternative anthropometric indicators of mortality. Am J Clin Nutr 1985; 42:296–306.
21. Briend A, Dykewicz C, Graven K, Mazumder RN, Wojtyniak B, Bennish M. Usefulness of nutritional indices and classifications in predicting death of malnourished

children. Br Med J (Clin Res Ed)1986; 293:373–5.
22. Briend A, Zimicki S. Validation of arm circumference as an indicator of risk of death in one to four year old children. Nutr Res 1986; 6:249–61.
23. Briend A, Wojtyniak B, Rowland MGM. Arm circumference and other factors in children at high risk of death in rural Bangladesh. Lancet 1987; 2(8561):725–8.
24. Chen LC, Chowdhury A, Huffman SL. Anthropometric assessment of energy-protein malnutrition and subsequent risk of mortality among preschool aged children. Am J Clin Nutr 1980; 33:1836–45.
25. Heywood PF. Nutritional status as a risk factor for mortality in children in the highlands of Papua New Guinea. Proceedings of the Thirteenth International Congress of Nutrition, Brighton, 1986:103–6.
26. Katz J, West KP Jr, Tarwotjo I, Sommer A. The importance of age in evaluating anthropometric indices for predicting mortality. Am J Epidemiol 1989; 130:1219–26.
27. Kielmann AA, McCord C. Weight-for-age as an index of risk of death in children. Lancet 1978; 1(8076):1247–50.
28. Smedman L, Sterky G, Mellander L, Wall S. Anthropometry and subsequent mortality in groups of children aged 6–59 months in Guinea-Bissau. Am J Clin Nutr 1987; 46:369–73.
29. Sommer A, Loewenstein MS. Nutritional status and mortality: a prospective validation of the QUAC stick. Am J Clin Nutr 1975; 28:287–92.
30. Vella V, Tomkins A, Ndiku J, Marshal T, Cortinovis I. Anthropometry as a predictor for mortality among Ugandan children, allowing for socio-economic variables. Eur J Clin Nutr 1994; 48:189–97.
31. Baker JP, Detsky AS, Wesson DE, Wolman SL, Stewart S, Whitewell J, Langer B, Jeejeebhov KN. Nutritional assessment: a comparison of clinical judgment and objective measurements. N Engl J Med 1982; 306:968–72.
32. Simoes EA, Desta T, Tessema T, Gerbresellassie T, Dagnew M, Gove S. Performance of health workers after training in integrated management of childhood illness in Gondar, Ethiopia. Bull World Health Organ 1997; 75(suppl 1):43–53.
33. Bern C, Zucker JR, Perkins BA, Otieno J, Oloo AJ, Yip R. Assessment of potential indicators for protein-energy malnutrition in the algorithm for integrated management of childhood illness. Bull World Health Organ 1997; 75(suppl 1):87–96.
34. Hamer C, Kvatum K, Jeffries D, Allen S. Detection of severe protein-energy malnutrition by nurses in The Gambia. Arch Dis Child 2004; 89:181–4.
35. Bairagi R. Effects of bias and random error in anthropometry and in age on estimation of malnutrition. Am J Epidemiol 1986; 123:185–91.
36. Velzeboer MI, Selwyn BJ, Sargent F, Pollitt E, Delgado H. The use of arm circumference in simplified screening for acute malnutrition by minimally trained health workers. J Trop Pediatr 1983; 29:159–66.
37. Davis LE. Epidemiology of famine in the Nigerian crisis: rapid evaluation of malnutrition by height and arm circumference in large populations. Am J Clin Nutr 1971; 24:358–64.
38. Loewenstein MS, Phillips JF. Evaluation of arm circumference measurement for determining nutritional status of children and its use in an acute epidemic of malnutrition: Owerri, Nigeria, following the Nigerian Civil War. Am J Clin Nutr 1973; 26:226–33.
39. Arnhold R. The QUAC stick: a field measure used by the Quaker Service team. J Trop Pediatr 1969; 15:243–7.
40. Shakir A, Morley D. Measuring malnutrition. Lancet 1974; 1(7860):758–9.
41. Shakir A. Arm circumference in the surveillance of protein-calorie malnutrition in Baghdad. Am J Clin Nutr 1975; 28:661–5.
42. Ross DA, Taylor N, Hayes R, McClean M. Measuring malnutrition in famines: are weight-for-height and arm circumference interchangeable? Int J Epidemiol 1990; 19:636–45.
43. English M, Esamai F, Wasunna A, Were F, Ogutu B, Wamae A, Snow RW, Peshu N. Assessment of inpatient paediatric care in first referral level hospitals in 13 districts in Kenya. Lancet 2004; 363(9245):1948–53.
44. Médecins Sans Frontières. Nutrition Guidelines. Paris: MSF, 1995.
45. Rivers JPW. The nutritional biology of famine. Biological Society Series. Oxford, UK: Oxford University Press, 1988:57–106.
46. Burgess HJ, Burgess AP. The arm circumference as a public health index of protein-calorie malnutrition of early childhood. (II) A modified standard for mid upper arm circumference in young children. J Trop Pediatr 1969; 15:189–93.
47. de Onis M, Yip R, Mei Z. The development of MUAC-for-age reference data recommended by a WHO Expert Committee. Bull World Health Organ 1997; 75:11–8.
48. Briend A, Garenne M, Maire B, Fontaine O, Dieng K. Nutritional status, age and survival: the muscle mass hypothesis. Eur J Clin Nutr 1989; 43:715–26.
49. Berkley J, Mwangi I, Griffiths K, Ahmed I, Mithwani S, English M, Newton C, Maitland K. Assessment of severe malnutrition among hospitalized children in rural Kenya: comparison of weight for height and mid upper arm circumference. JAMA 2005; 294:591–7.
50. Berkley J, Newton C, Maitland K. Severe malnutrition assessment in children in rural Kenya. JAMA 2005; 294:2577.
51. Waterlow JC. Classification and definition of protein-calorie malnutrition. Br Med J 1972; 3:566–9.
52. Spector J. Community-based therapeutic care (CTC) considerations in South Wollo, Ethiopia, including an assessment of MUAC and height indices as the sole screening and admission criteria for entry into the outpatient therapeutic program (OTP) (internal report). Oxford, UK: Valid International, 2005.
53. Trowbridge FL, Sommer A. Nutritional anthropometry and mortality risk. Am J Clin Nutr 1981; 34:2591–2.
54. Kahigwa E, Schellenberg D, Schellenberg JA, Aponte JJ, Alonso PL, Menendez C. Inter-observer variation in the assessment of clinical signs in sick Tanzanian children. Trans R Soc Trop Med Hyg 2002; 96:162–6.
55. Sandiford P, Paulin FH. Use of mid-upper-arm circumference for nutritional screening of refugees. Lancet 1995; 345(8957):1120.
56. Pelletier DL, Frongillo EA, Habicht JP. Epidemiologic evidence for a potentiating effect of malnutrition on child mortality. Am J Public Health 1983; 83:1130–3.
57. Briend A, Golden MHN, Grellety Y, Prudhon C, Hailey P. Use of mid-upper-arm circumference for nutritional screening of refugees. Lancet 1995; 345(8957):1119–20.