**Technical Background Paper**

# 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[1], Tanya Khara[2] and Steve Collins[2]

[1] Division of Epidemiology, University College London, Institute of Ophthalmology, 11-43 Bath Street, London, EC1 V 9EL, UK
[2] Valid International, Valid International Ltd, Unit 14 Oxford Enterprise Centre, Standingford House, 26 Cave Street, Oxford, OX4 1BA, UK

Address for correspondence:
Mark Myatt – mark@brixtonhealth.com
Alltgoch Uchaf
Llawryglyn
Caersws
Powys SY17 5 RJ
Wales
UK



This paper was written as a Technical Background Paper for an Informal Consultation held in Geneva (21-23 November 2005) to discuss the community-based management of severe malnutrition. The meeting was organised by the Department of Child and Adolescent Health and Development and the Department of Nutrition for Health and Development of the World Health Organization, UNICEF and the Standing Committee on Nutrition of the United Nations.

## Abstract

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 using MUAC but being subsequently refused treatment because they do not meet the weight-for-height admission criteria. This "problem of rejected referrals" has proved to be a major barrier to program uptake. 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. MUAC with the addition of the presence of bipedal oedema was found to be the indicator best suited to screening and case-detection for malnutrition in the community. The case-definition "MUAC < 110 mm OR the presence of bipedal oedema" with MUAC measured using a colour banded strap is suitable for screening and case-detection for malnutrition in the community for children aged between six and fifty-nine months. Monitoring and discharge criteria were also reviewed and it was found that there was no compelling evidence to support a move away from using weight in combination with clinical criteria for monitoring and discharge.

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

## 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 their family (e.g. risk of nosocomial infection, loss of carer for siblings, loss of labour 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, called community-based therapeutic care (CTC), designed to address the limitations of inpatient care has been proposed [3]. CTC programs use decentralised networks of outpatient treatment sites (usually located at existing primary healthcare 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 environment. Persons with severe malnutrition, good appetite, and without medical complications are treated in an *Outpatient Therapeutic Program* (OTP) that provides *ready-to-use therapeutic food* (RUTF) and medicine to treat simple medical conditions. These are taken at home and the patient attends an OTP site weekly or fortnightly for monitoring and re-supply. Severely malnourished persons with medical complications and / or anorexia are treated in an inpatient *Stabilisation Centre* (SC) where they receive standard WHO initial care until they have appetite and are well enough to continue with outpatient care [4]. CTC programmes have treated over 9000 severely malnourished children in Ethiopia, Malawi, and Sudan, meeting SPHERE 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 with a weight-for-height (W/H) z-score below -3.00 or a weight below 70% of the median W/H using the NCHS reference population median and / or bipedal oedema 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 using mid-upper-arm-circumference (MUAC), measured in the community by community-based volunteers, for referral and W/H, measured at the treatment site by program staff, for admission 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 this *problem of rejected referrals* leads to carers of referred children becoming unwilling to attend for admission even when their children's condition deteriorates, carers of rejected children actively disparaging the program, local leaders becoming disillusioned with the program, as well as low levels of staff and volunteer morale and performance [7, 8, 6]. In some programs the problem of rejected referrals was solved by moving towards a unified MUAC-based referral and admission criteria [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].

The large numbers of children referred to treatment sites for second-stage screening by a two-stage referral and admission system also tend to lead to crowding and long waits at treatment sites and the diversion of often scarce resources away from treatment and carer-education towards crowd-control and second-stage screening activities. Long waits at treatment centres negatively impact upon the community's perception of programs and this negatively impacts upon program coverage [11, 6]. 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 where 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 healthcare 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 methodolical framework

The defining characteristics of an appropriate case-detection method depends 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 non-clinicians.

**ACCEPTABILITY :** The method is acceptable to the subject and others.

**COST :** The overall cost of the method.

**PRECISION :** The degree of reproducibility amongst 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.

They identify four distinct contexts in which case-detection methods are applied: Epidemiological 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 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 factors relating to program organisation 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 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 where 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].

**Table 1 :** Relative importance of key properties of case-detection methods in different contexts

| Property | Context | | | |
|---|---|---|---|---|
| | **Epidemiological survey / surveillance** | **Screening and case-detection in the community** | **Case-finding in clinical contexts** | **Diagnosis in clinical contexts** |
| **Simplicity** | ++++ | ++++ | – | – |
| **Acceptability** | ++++ | +++ | + | – |
| **Cost** | ++++ | ++ | – | – |
| **Objective** | ++++ | ++++ | ++++ | ++++ |
| **Quantitative** | ++++ | ++++ | – | – |
| **Independent of age** | ++++ | ++++ | – | – |
| **Precision (reliability)** | + (individual) ++++ (group) | ++ | ++++ | ++++ |
| **Accuracy** | + (individual) ++++ (group) | ++ | ++++ | ++++ |
| **Sensitivity** | + | ++ | +++ | +++ |
| **Specificity** | + | ++++ | ++++ | ++++ |
| **Predictive value** | + | ++ | ++++ | ++++ |

– irrelevant / + minor / ++ moderate / +++ major / ++++ crucial

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 non-clinicians” 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 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 (e.g.) a threshold value of an anthropometric indicator of nutritional status, a large proportion of deaths in untreated individuals (i.e. 50% or more) should occur in children below the case-defining threshold. Deaths occurring below the case-defining threshold are likely to be related to nutritional status and to respond to dietary treatment. Case-detection method 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* have been reversed from those 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 favours 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 characterised by extreme scarcity of capacity relative to need and in which a false positive misdiagnosis may have negative consequences for the subject and their family as well as high financial cost to the provider. It may not be well suited to alternative models of delivery, such as the Community Therapeutic Care (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 minimise 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 utilisation 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 categorised 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 oedema.
>
> **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 specimen collection, specimen storage, specimen transit, and the reporting of 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 compared to a threshold value. Individuals in which 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 criteria 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 would be likely to survive if treated in an appropriate nutritional support program. This realisation has led a number of workers to argue that the utility of case-definitions for malnutrition are defined more by their ability to reflect mortality risk than than their ability to reflect nutritional status [18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 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 to 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] report good agreement between the clinical diagnosis of malnutrition made by trained nurses and by a reference paediatrician in primary care settings in Ethiopia. Bern et al. [33] also report good results with a single trained health worker in a district hospital in Kenya using visible severe wasting and / or bipedal oedema as the case-definition for severe malnutrition. This finding is, however, problematic as 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 calculated at the time of the clinical assessment by the same health worker who performed the clinical assessment. Hamer et al. [34] report poor results using the same case-definition and validation criteria with trained registered and auxiliary nurses in a tertiary level referral hospital in The 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] reports that indicators that include an age component (i.e. H/A, W/A, 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 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] report that minimally trained community health volunteers in rural Guatemala had difficulties performing date arithmetic.

Multi-component indicators (i.e. W/A, H/A, W/H, MUAC/A, MUAC/H) usually require finding values by look-up in multi-dimensional 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 (i.e. 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] test the post-training performance of five minimally trained community health volunteers in rural Guatemala on their ability to calculate the W/H indicator, and report that four of the five could not complete the test unsupervised due to problems with rounding decimal numbers (required for table look-up) and that the one worker who completed the test unsupervised required over an hour to calculate ten indicator values of which four were incorrect. Hamer et al. [34] report that registered and auxiliary nurses in a tertiary level referral hospital in The Gambia had difficulties using growth charts immediately after training. It is unlikely, therefore, that these tasks could be performed by minimally trained community-based volunteers.

Sommer and Lowenstein [29] report that MUAC/H, when measured with a device known as a QUAC stick, is a multi-component indicator that does not require table look-up or reference to a “growth chart”. The QUAC stick avoids table look-up by having the MUAC thresholds defining malnutrition marked on a “height” stick. A child taller than the corresponding mark on the “height” stick for their 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] reports 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 Lowenstein 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 report 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, table look-up, or plotting of data on “growth charts”. Shakir and Morley [40] suggest the use of a colour-banded cord to measure MUAC, with colours corresponding to classifications of malnutrition. Shakir [41] reports that a colour-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 using a readily understandable “traffic light” system intuitively related to thinness may have a potential for raising awareness amongst community members of the prevalence of malnutrition which is an essential first step in the process of mobilising community action to counter the problem.

## Acceptability

Velzeboer et al. [36] in a comparison of W/H and MUAC in Guatemala report that younger children tended to become upset and agitated during both weight and height measurement and that no such behaviour was observed during the measurement of MUAC. Their characterisation of these children as “traumatised” may be a little strong as any trauma is unlikely to have lasting consequences. The unpleasantness associated with weight and height measurement may, however, reduce the acceptability of indicators using weight and / or height measurements to children, their carers, and community-based volunteers and negatively impact upon the coverage of case-detection activities particularly if carers of sick children refuse to have their child 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 using 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 centralised programs with dedicated case-detection teams but is likely to prove unacceptable in programs relying on decentralised networks consisting of large numbers of community-based volunteers for case-detection. MUAC and MUAC/H measured using the QUAC stick can be performed using low-cost and maintenance-free equipment [37, 40, 41]. Weight and height measurements are generally considered to require three persons (two to measure, one to supervise, record, and calculate indicator values) to collect with precision and accuracy [44]. It may prove difficult to find a sufficient number of sufficiently 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.

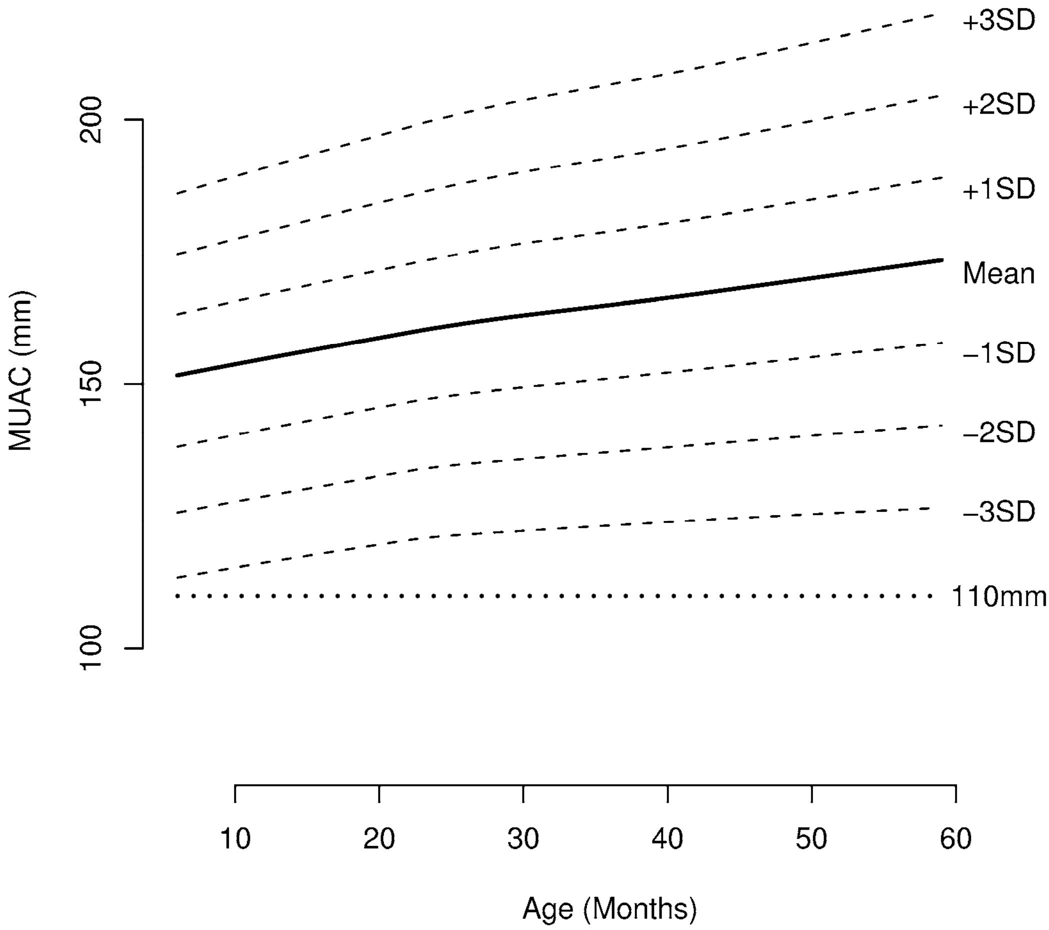
## Objective and quantitative

The subjective nature of clinical assessment may lead to acceptability problems as carers may feel that non-clinical criteria (i.e. social, racial, or tribal discrimination) are being applied. Corruption is also an issue that must be considered with any subjective criteria. Clinical assessment is generally recognised as subjective, difficult to standardise, and difficult to express quantitatively [16, 37, 34]. 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. 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 dates of birth or exact ages are unknown this is likely to be a major problem. Children grow fast and this means that small errors in estimating age may lead to large errors in indicator values. In famine, and situations where displacement and familial separation are common, field workers are often required to estimate the age of children based on little or no information. Estimates "by eye" are biased by assumptions about the relationship between height and age which 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 one and five years [16, 19, 30, 37, 42, 46, 47] but is age-dependent in children aged below one year [47]. The relationship between MUAC and age is shown in *Figure 1*. The predictive power (i.e. of predicting mortality) of MUAC is, however, independent of age even in children aged below one year [22, 30, 48, 49, 50]. Berkley et al. [50] report that consistently high case fatality rates in hospitalised Kenyan children of all ages between 12 and 59 months with low MUAC values, which they define as less than or equal to 115 mm, suggesting that unadjusted (i.e. by age) MUAC may be useful in clinical settings. W/H is also independent of age between the ages of one and five years [42, 51] but the predictive power (i.e. of predicting mortality) of W/H may change with age [26].

**Figure 1 :** MUAC/A growth reference curves for males and females aged between 6 and 59 months




The MUAC/A growth reference curves presented in this figure are taken from de Onis et al. (1997) [47]

## 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 includes an age component [35, 45]. It is often asserted that, in terms of precision and accuracy of measurement, MUAC compares unfavourably 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. This assumption has not been tested in children but has been demonstrated to be invalid in adults in famine situations and in labour 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 gastro-intestinal tract and that heavy parasitism with *Ascaris lumbricoides* may bias weight upwards. Davis [37] reports that MUAC/H measured using a QUAC stick was both reproducible and accurate. This finding is confirmed by Sommer and Lowenstein [29]. Velzeboer et al. [36] test the reliability (i.e. precision) of five minimally trained community health volunteers in rural Guatemala for W/H, H/A, W/A, MUAC, and MUAC/A and report 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 report that under field conditions, minimally trained workers made fewer and smaller errors with MUAC than with W/A or W/H, even when not required to calculate indicator values by table look-up or by plotting data on “growth charts”.

Feeney [9] reports 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 (i.e. using Roman rather than Amharic numerals). Recording errors did not have operational consequences since referral for admission was informed by the subject's position with regard to a threshold value. A companion study found that when asked to classify children according to whether or not their MUAC fell below a fixed threshold of 110 mm, very few errors were made [9]. Feeney [9] and Spector [52] both identify 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

Lowenstein and Phillips [38] and Sommer and Lowenstein [29] report MUAC/H as strongly predictive of death at one, three, and eighteen months after measurement. Kielman and McCord [27] report that W/A was predictive of death at six and twelve months after measurement in Indian children. Chen et al. [24] examine 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. Sommer [53], analysing a subset of the data reported by Chen et al. [24], reports that MUAC alone performed better than MUAC/H and that adjusting MUAC 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 Lowenstein [29] in a study to validate the use of MUAC as an indicator of risk of death within one, three, and six months of measurement in Bangladeshi children, report that MUAC alone performed better in terms of both sensitivity and specificity compared with 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 and this 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 centralised screening teams [18]. Briend et al. [22] examine the predictive power of W/A, W/H, H/A, MUAC, and MUAC/A for predicting death in children hospitalised with diarrhoea in a Dhaka hospital and report 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 may have higher predictive power and found no combination of indicators that outperformed MUAC alone. Briend et al. [23] report that MUAC, as an indicator of risk of death within one month of measurement in Bangladeshi children, was almost twice as sensitive at the same specificity than other anthropometric indicators and that only slight improvements in sensitivity could be achieved using a diagnostic algorithm using MUAC and selected clinical signs. Alam et al. [19] examining MUAC, MUAC/A, MUAC/H, H/A, W/H, and H/A for predicting death three and six months after measurement in Bangladeshi children report 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] report 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] report H/A but not W/H as a significant predictor of mortality in Bangladeshi children. Vella et al. [30] test the predictive power of W/A, H/A, W/H, and MUAC in Ugandan children and found that MUAC was the most sensitive, in relation to specificity, predictor of mortality within twelve 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 whilst other indicators did not improve the predictive power of MUAC. Berkley et al. [49] report 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 in some contexts *marasmus* is the predominant form of severe malnutrition whereas in others *kwashiorkor* is the predominant form of severe malnutrition [16]. This problem is usually addressed by using an anthropometric indicator to define marasmus and the presence or absence of bipedal oedema to define kwashiorkor [51]. Kahigwa et al. [54] report substantial agreement between two clinical officers in a Tanzanian hospital for identification of oedema. Hamer et al. [34] report that trained registered and auxiliary nurses in a tertiary level referral hospital in The Gambia performed poorly at identifying bipedal oedema and it was observed that the nurses spent insufficient time depressing tissues. Simoes et al. [32] report good agreement between the clinical diagnosis of malnutrition made by trained nurses and by a reference paediatrician 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 oedema) 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] report that MUAC used alone was more sensitive and more specific than either W/H and W/A used alone as a test for bipedal oedema in Malawi. Berkley et al. [49] report that MUAC used alone performed better than W/H used alone at identifying children with bipedal oedema and skin and hair changes associated with kwashiorkor in Kenya. Currently available data suggests that the use of MUAC may, to some extent, compensate for the potentially poor performance of minimally trained community-based volunteers at identifying bipedal oedema by clinical examination.

## The use of anthropometry in young children

Anthropometric measurements are difficult to perform on young children. Children under six months of age weigh only a few kilograms. To obtain sufficiently accurate measurements of weight, children aged less than six months should be weighed using specialist paediatric scales which are graduated in units of 10g rather than conventional hanging scales which are graduated in units of 100g. This requires the provision and maintenance of suitable scales. Measuring the length of children aged less that six months can be done 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. Also, internationally recognised 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. Weight measurement in adolescents requires physician scales. Height measurement in adolescents requires height boards capable of measuring heights of two metres 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

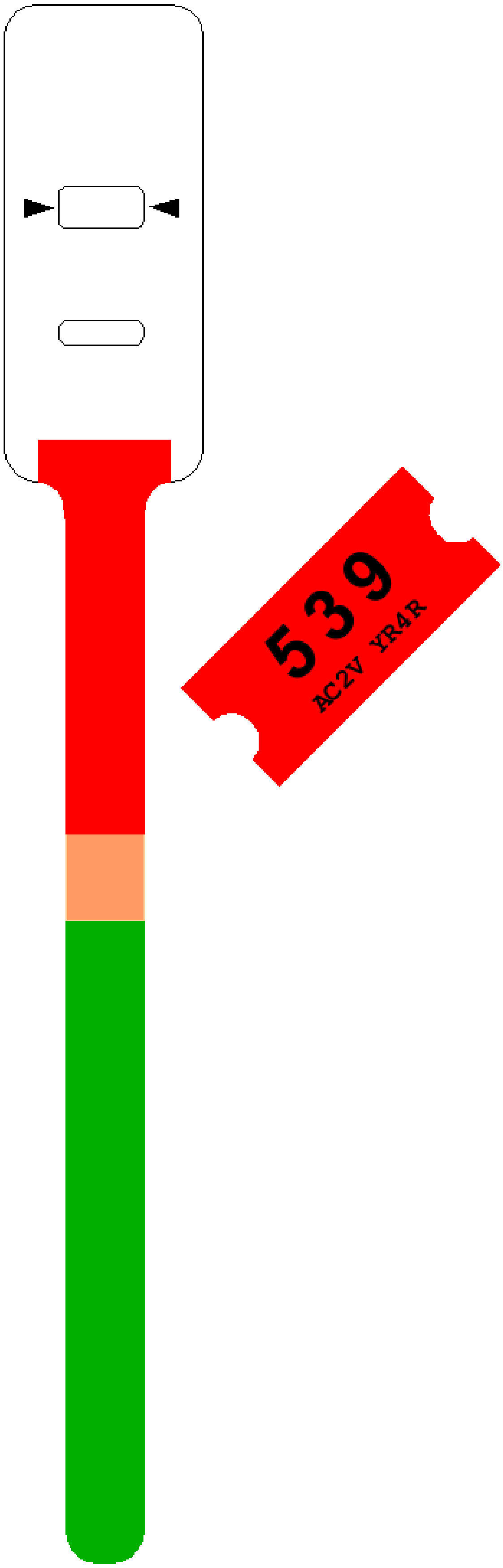
*Table 2* summarises the data presented above according to whether specific indicators exhibit the key properties outlined in the conceptual and methodological framework. Within this framework, MUAC or MUAC/H measured using the QUAC stick with the addition of the presence of bipedal oedema are the indicators 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 methodological framework used in this report, one of the least useful indicators to use in this context.

**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

| **Property** | **Indicator** | | | | | | |
|---|---|---|---|---|---|---|---|
| | **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) |
| **Objective** | No | No | No | Yes | Yes | No | Yes |
| **Quantitative** | No | Yes | Yes | Yes | Yes | Yes | Yes |
| **Independent 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 |

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 tickets” or “raffle tickets”. Re-measurement 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 (*Figure 2*).

**Figure 2 :** Banded MUAC strap and cloakroom / raffle ticket referral slip

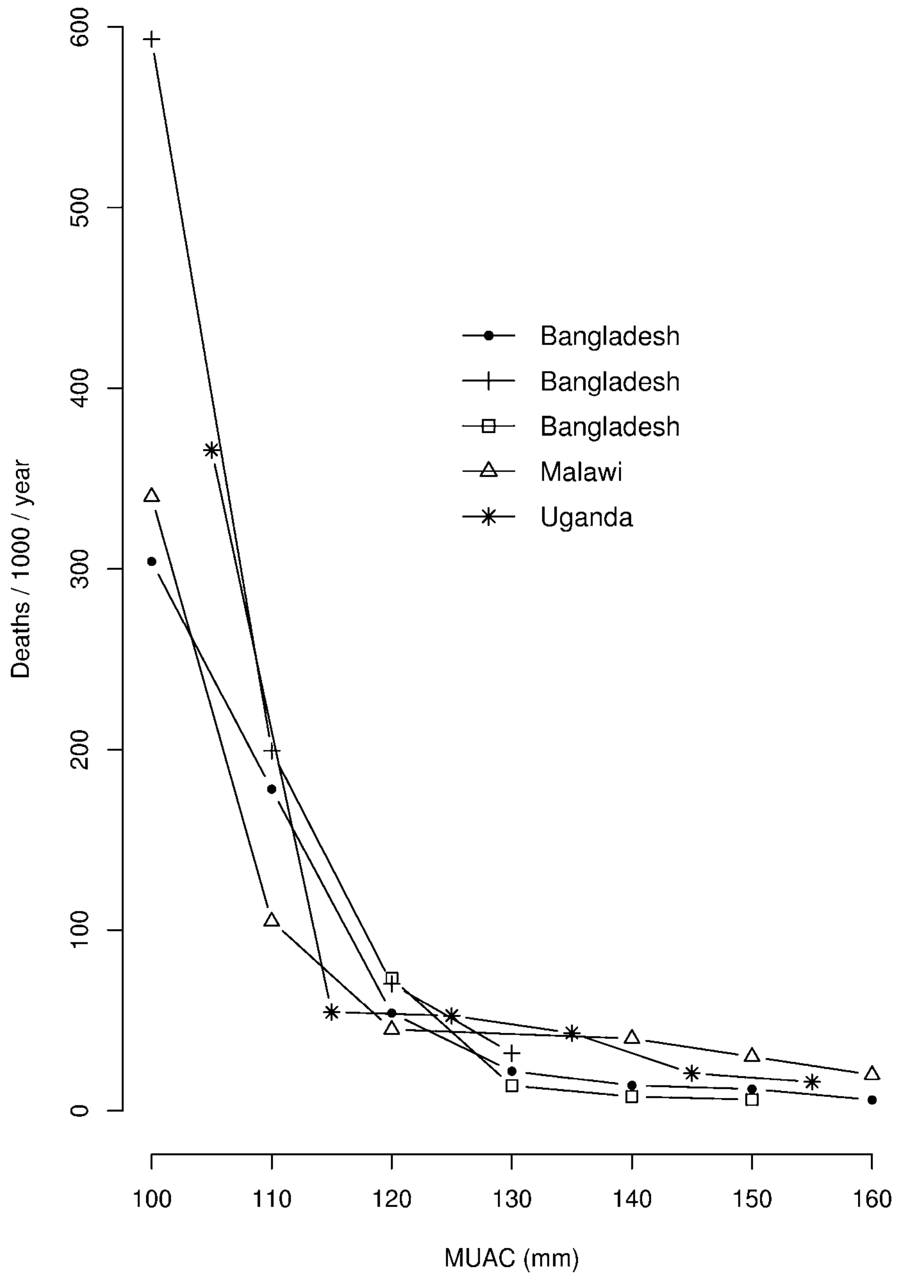


## Selecting an appropriate indicator threshold

Using an anthropometric indicator such as MUAC in case-definitions of malnutrition requires that the indicator is measured and the value of the indicator compared to a threshold value. Individuals in which 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 colour-coded on the strap providing a simple-to-use, instant, 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 1000 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 mm 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 suggests 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 three z-scores below the mean of the sex-combined MUAC/A reference distribution for children aged seven months or older and equal to or more extreme than four z-scores below the mean of the sex-combined MUAC/A reference distribution for children aged 39 months or older [47].

**Figure 3 :** Observed relationship between MUAC and child mortality in five studies



## A proposed case-definition

Currently available data suggests that the case-definition:

MUAC < 110 mm OR the presence of bipedal oedema

with MUAC measured using colour-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 for children aged between six months and five years. Height may be used as a proxy for age. In this case the proposed case-definition applies only for children between 65 cm and 110 cm in height with eligibility ascertained using a simple marked stick. These height thresholds are conventional and may not be appropriate in setting 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 utilisation

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 indicates that MUAC is one of the best predictors of mortality, but children selected for treatment because they have extreme 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*.

**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 |

The intensity of intervention that is required for children with extreme low values of MUAC is also of interest. If children with extreme 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 outpatient therapeutic program (OTP), would be an inappropriate use of resources. This question is of particular interest in smaller children, usually defined as age less than twelve months or height less than or equal to 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:

Do smaller children with extreme low values of MUAC do well in OTP?

Do smaller children with extreme low values of MUAC do well in SFP?

A *natural experiment* in a CTC program in Northern Ethiopia in 2003 provides answers to these questions for smaller children without bipedal oedema 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 following case-definition:

MUAC $<$ 110 mm AND (age $>$ 12 months OR height $>$ 75 cm) AND W/H $>$ 70%

were admitted to the outpatient therapeutic program (OTP).

In March 2003 this 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, amongst children with a MUAC below 110mm, the smaller ones (i.e. those whose height was less than or equal to 75 cm) from admission to the OTP. This change in case-definitions created a natural experiment consisting of two comparable groups of children with MUAC below 110 mm, height less than or equal to 75 cm, W/H greater than 70% of the reference median, and without bipedal oedema being admitted initially to OTP and then to SFP. This was noted during a program review in November 2003 and allowed a comparison of response of smaller children with extreme low values of MUAC admitted to OTP and SFP. Summary data from the natural experiment are presented in *Table 4*.

**Table 4 :** Summary of data arising from a natural experiment allowing comparison of response to treatment of children with MUAC $<$ 110 mm, height $\leq$ 75 cm, W/H $>$ 70% of the reference median, and without oedema in OTP and SFP

| | **Experimental arm** | |
|---|---|---|
| | **OTP** | **SFP** |
| **Number of subjects** | 42 | 56 |
| **Survived** | 40 | 46 |
| **Died** | 0 | 8 |
| **Lost to follow-up / defaulted** | 2 | 2 |
| **Age range (median)** | 12 – 36 months (16 months) | 6 – 36 months (14 months) |
| **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** | 54% male | 57% male |

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 suggest preferential reporting of age as thirteen 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 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 adjusting for age at admission split into age < 13 months and age ≥ 13 months (Mantel-Haenszel $\chi^2$ = 3.86, df = 1, p = 0.0494). This analysis is compromised by probable inaccurate reporting and / or recording of age. The effect observed in the crude analysis remains statistically significant after adjusting for height (i.e. as a proxy for age) at admission split into height above or below the overall median height at admission of 66.15 cm (Mantel-Haenszel $\chi^2$ = 4.89, df = 1, p = 0.0269).

**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 oedema in OTP and SFP

| | | Outcome ... | | |
|---|---|---|---|---|
| | | Died | Survived | |
| Exposure : | SFP | 8 | 46 | 54 |
| | OTP | 0 | 40 | 40 |
| | | 8 | 86 | 94 |

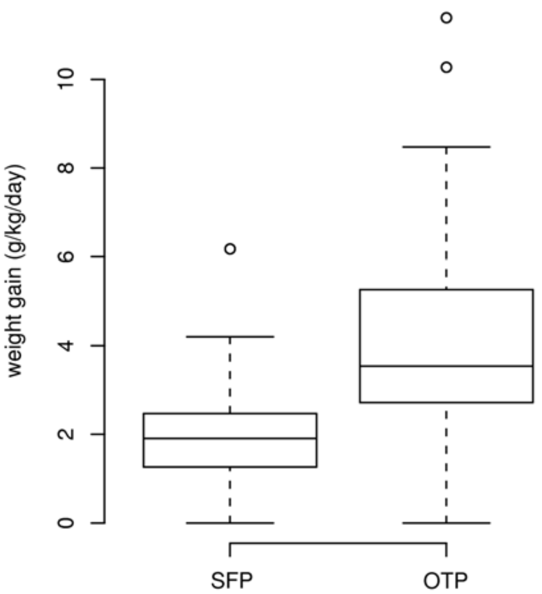
```
Fisher-Irwin exact test : p = 0.0094 (one sided)
                          p = 0.0191 (two-sided)

        Risk Difference : 14.81%, 95% CI = 3.15% - 26.47%
                 z-test : z = 2.17
                          p = 0.0149 (one-sided)
                          p = 0.0299 (two sided)
```

*Figure 4* shows the results of an analysis of weight gains in g/kg/day observed in the two arms of the natural experiment. Smaller children with MUAC less than 110 mm responded well (i.e. in terms of both survival and weight gain) to the high intensity intervention (i.e. OTP) but did not respond well to the low intensity intervention (i.e. SFP). Treating such children with a high intensity intervention such as therapeutic feeding in an outpatient therapeutic program (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 less than or equal to 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 have received 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 seven months with two 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.

**Figure 4 :** Observed weight gains (g/kg/day) 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 oedema in OTP and SFP




```
--- SFP ---
                   Sample size : 56
    Mean weight gain (g/kg/day) : 1.89
            Standard deviation : 1.13
  Median weight gain (g/kg/day) : 1.90

--- OTP ---
                   Sample size : 42
    Mean weight gain (g/kg/day) : 4.28
            Standard deviation : 2.44
  Median weight gain (g/kg/day) : 3.55

--- Difference ---
      Kruskal-Wallis Chi-square : 34.6714
                             df : 1
                              p : 0.0000
```

## 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 oedema

Changing this to:

MUAC < 110 mm OR the presence of bipedal oedema 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-definition. This was tested using a simple computer-based simulation. More than 200 datasets from nutritional anthropometry surveys that collected data on sex, weight, height, MUAC, and oedema were collected from international NGOs. These datasets were restructured to ensure compatible coding between datasets and combined into a single large dataset representing over two hundred and ten thousand children between 65 cm and 110 cm in height. The prevalence of malnutrition in the combined dataset using standard W/H-based case-definitions is summarised in *Table 6*.

**Table 6 :** Prevalence of malnutrition in the combined dataset using case-definitions based on W/H z-scores and the presence / absence of bipedal oedema

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

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

**W/H-based case-definition :**

(MUAC < 125 mm AND W/H < 70% of the reference median) OR OEDEMA

**MUAC-based case-definition :**

MUAC < 110 mm OR OEDEMA

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:

Malnourished children identified by the MUAC-based case-definition : 5484

Malnourished children identified by the W/H-based case-definition : 3678

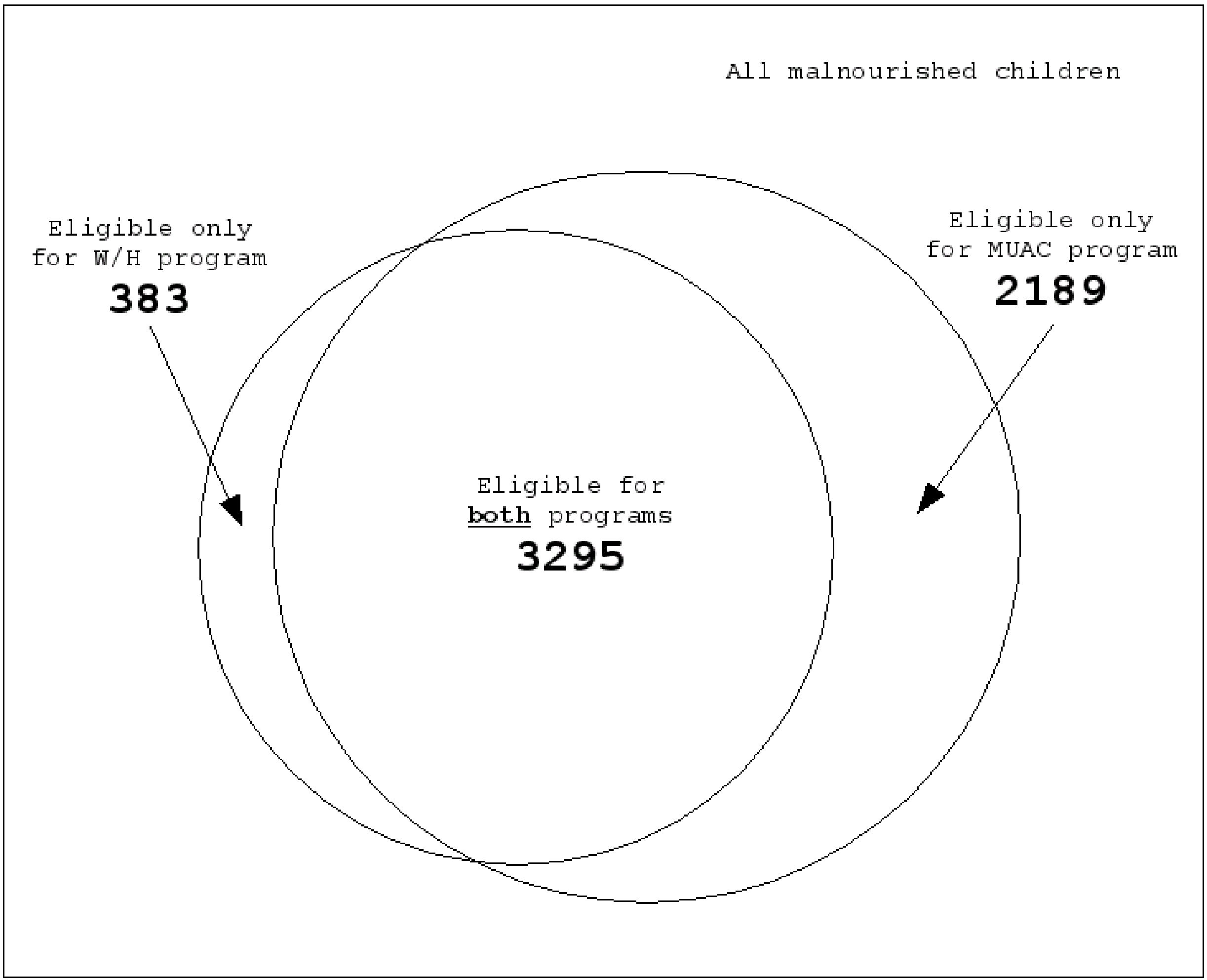
Overall need in the combined dataset was estimated as the number of children identified as severely malnourished by either case-definition [57].

The MUAC-based and W/H-based case-definitions selected many of the same children. When using the MUAC-based case-definition, the number of excluded low W/H children is small relative to estimated need:

Estimated overall need (i.e. children identified as malnourished by either the MUAC-based or the W/H-based case-definition) : 5867

Malnourished children excluded by the MUAC-based case-definition : 383 (6.53%)

**Figure 5 :** Number of children selected from the combined dataset using two different case-definitions†




†The two case-definitions used are:

**W/H program :** (MUAC < 125 mm AND W/H < 70% of the reference median) OR OEDEMA

**MUAC program :** MUAC < 110 mm OR OEDEMA

Overall need (labelled as "All malnourished children") is the number of children identified by either case-definition.

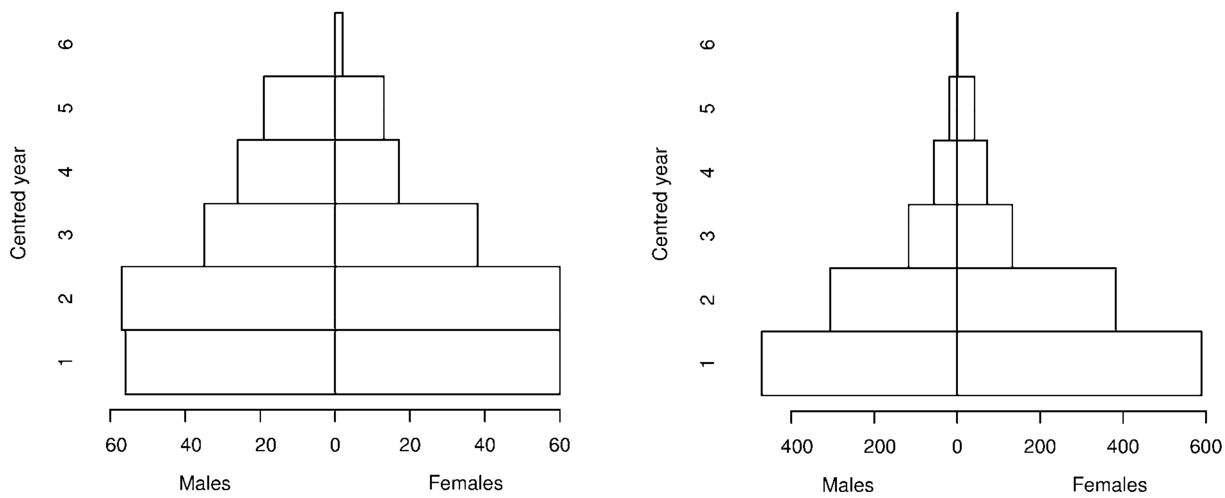
When using W/H-based case-definition, however, the number of excluded low MUAC children is large relative to estimated need:

> Estimated overall need (i.e. children identified as malnourished by either the MUAC-based or the W/H-based case-definition) : 5867
>
> Malnourished children excluded by the W/H-based case-definition : 2189 (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.

**Figure 6 :** Age-profiles of children excluded using two different case-definitions.

Included by W/H but **excluded by MUAC**     Included by MUAC but **excluded by W/H**




Sample size : 383
Median age (months) : 24

Sample size : 2189
Median age (months) : 18

Kruskal-Wallis Chi-square = 55.8341, df = 1, p = 0.0000

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 (i.e. as a result of simplicity, acceptability, low cost, and effective use of community-based volunteers and program staff) than case-finding activities using a

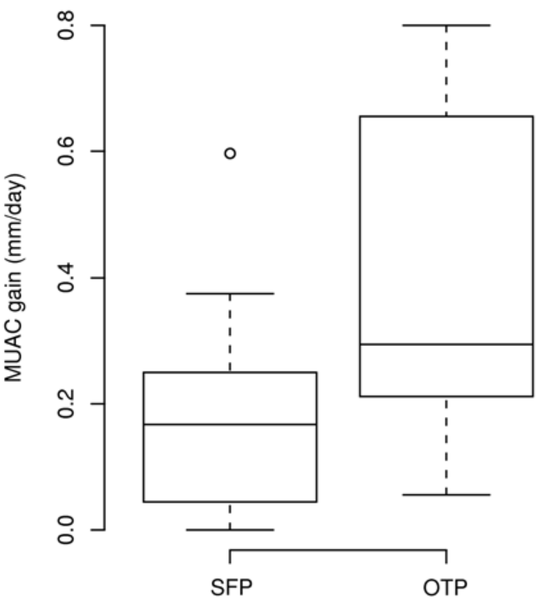
W/H-based case-definition. Programs using a MUAC-based case-definition are likely to have a higher uptake (i.e. as a result of minimising the problems of rejected referrals, crowding, and long waiting times) than programs using a W/H-based case-definition. The results presented in *Figure 5* are, therefore, 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 epidemiological and needs-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 demonstrates 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 [57].

**Figure 7 :** Observed MUAC gains (mm/day) 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 oedema in OTP and SFP




```
--- SFP ---
            Sample size : 24
 Mean MUAC gain (mm/day) : 0.17
     Standard deviation : 0.16
Median MUAC gain (mm/day) : 0.15

--- OTP ---
            Sample size : 19
 Mean MUAC gain (mm/day) : 0.40
     Standard deviation : 0.25
Median MUAC gain (mm/day) : 0.30

--- Difference ---
Kruskal-Wallis Chi-square : 8.9913
                     df : 1
                      p : 0.0027
```

At present there is no compelling data to suggest a move away from a weight-based indicator towards a MUAC-based indicator for monitoring and discharge. It should be noted, however, that height boards are often unavailable in primary healthcare centres 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 MUAC who are already above the W/H discharge criteria. Current practice in CTC programs for such cases is to monitor weight and:

> Discharge as *cured* after a minimum of two months in OTP if MUAC > 110 mm, no oedema for a minimum of two weeks, evidence of sustained weight gain, and patient is “clinically good”.
>
> Discharge as a *non-responder* after a minimum of four months in OTP if weight is stable and all available treatment options (e.g. home visits, inpatient stabilisation, hospitalisation, ART programs, TB treatment programs) have been pursued.

These monitoring and discharge criteria may be applied to all cases. The advantage of this approach is that it only requires that weight is monitored and suitable scales are usually available in primary healthcare centres in developing countries operating growth monitoring programs. Monitoring weight alone does not differ greatly from monitoring W/H since height changes little during recovery and changes in W/H are mainly due to changes in weight rather than height and that, 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 is 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 criteria. With this approach patients would be discharged once their percentage weight gain exceeded a cut-off value based on their weight at admission (or weight at loss of oedema for patients presenting with *marasmic kwashiorkor*). Preliminary analysis of data from CTC programs in Malawi and Ethiopia suggests that a cut-off of 15% would result in c. 50% of discharges meeting or exceeding 80% of the W/H reference median and that a cut-off of 18% would result in c. 50% of discharges meeting or exceeding 85% of the W/H reference median. Percentage weight gain could be combined with a MUAC cut-off. 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 look-up table. For example, *Table 7* shows discharge weights for admission weights based on a 15% weight gain.

**Table 7 :** Example look-up table for calculating percentage change (15% in this table) in weight

| 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 |

There are aspects of CTC programs (e.g. the concentration on maximising program coverage and community-based delivery of services) that are more typical of "public health" or "mass treatment" interventions than traditional centre-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 oedema for admission. In such programs, patients admitted with oedema but with a W/H percentage of median above 80% are, typically, retained in the program for a fixed period after loss of oedema. Preliminary analysis of data from CTC programs in Malawi and Ethiopia suggests that an episode length of 60 days would result in c. 50% of discharges achieving ≥ 15% weight gain at discharge.

As data from CTC programs becomes 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, or 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 towards a 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 oedema) for monitoring and discharge.

# References


1. Van Damme W, Medical assistance to self-settled refugees. Guinea 1990-1996, ITG Press, Antwerp, 1998:1-249

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;358:498-501.

4. World Health Organization, Management of severe malnutrition: A manual for physicians and other senior health workers (1st Ed.), WHO, Geneva, 1999

5. SPHERE Project, The SPHERE Humanitarian Charter and Minimum Standards in Disaster Response (2nd Ed., Summer 2003 Draft), The SPHERE Project, Geneva, 2003

6. Collins S, Community-based therapeutic care : A new paradigm for selective feeding in nutritional crises (Network Paper No. 48), Humanitarian Practice Network, 2004

7. Khara, T, Report on focus group discussions: Investigation into beneficiary opinions of the CTC programme at distribution sites, INTERNAL REPORT, VALID International / CONCERN Worldwide, 2003

8. Golden K, Khara T, CTC in Ethiopia: Working from CTC principles, In: Collins S, Khara T, Community-based Therapeutic Care (CTC), Emergency Nutrition Network (ENN) Special Supplement. No.2, 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, London School of Hygiene and Tropical Medicine (MSc Thesis), 2004

10. CONCERN, MSF(CH), SCF(US), Community-based therapeutic care (CTC) in Ethiopia: Proceedings of a workshop held in Addis Ababa, Ethiopia, 22nd-23rd June 2004, UNPUBLISHED REPORT, 2004

11. Khara T, Collins S, Emergency Nutrition Network Special Supplement : Community Based Therapeutic Care, Emergency Nutrition Network Supplement Series No. 2, 2004

12. Lee J, Improving the Community-based Therapeutic Care programme coverage, Dowa district, Malawi. INTERNAL REPORT, 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, VALID International, 2003

14. Sackett DL, Holland WW, Controversy in the detection of disease, Lancet, 1975;25:357-359

15. Beaton GH, Bengoa JM, Practical population indicators of health and nutrition, In: Nutrition and Preventive Medicine, G Beaton and JM Bengoa (Eds.), Monograph Series 62, World Health Organisation, 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: 179-192.

17. Habicht JP, Some characteristics of indicators of nutritional status for use in screening and surveillance, Am J Clin Nutr, 1980;33:531-535

18. Pelletier DL, The relationship between child anthropometry and mortality in developing countries: Implications for policy, programs, and future research, J Nutr, 1994;124:2047S-2081S

19. Alam N, Wojtyniak B, Rahaman MM, Anthropometric indicators and risk of death, Am J Clin Nutr, 1989;49:884-888

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, Dykewixz C, Graven K, Mazunder RN, Wojtyniak B, Bennish M, Usefulness of nutritional indices and classifications in predicting the death of malnourished children, Br Med J, 1986;293:373-275

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;6:249-261

23. Briend A, Wojtyniak B, Rowland MGM, Arm circumference and other factors in children at height risk of death in rural Bangladesh, Lancet, 1987;26:725-727

24. Chen LC, Chowdhury MK, Huffman SL, Anthropometric assessment of energy-protein malnutrition and subsequent risk of mortality among pre-school children, Am J Clin Nutr, 33; 1980;33:1836-1845

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, 1986:103-106

26. Katz J, West KP, Tarwotjo I, Sommer A, The importance of age in evaluating anthropometric indices for predicting mortality, Am J Epidemiol, 1989;130(6):1219-1226

27. Kielman AA, McCord C, Weight-for-age as an index of risk of death in children, Lancet, 1978;10:1247-1250

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;46:396-73

29. Sommer A, Lowenstein MS, Nutritional status and mortality: A prospective validation of the QUAC stick, Am J Clin Nutr, 1975;28:287-292

30. Vella V, Tomkins A, Ndiku J, Marshal T, Cortinovis I, Anthropometry as a predictor for mortality among Ugandan children allowing for socio-economic status, Eur J Clin Nutr, 1994;48:189–197

31. Baker JP, Detsky, AS, Wesson DE, Wolman SL, Stewart S, Whitewell J, Langer B, Jeejeebhov KN, Nutritional assessment: A comparison or clinical judgement and objective measurements, N Engl J Med, 1982;306(16):968-972

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;75:87-96

34. Hamer C, Kvatum K, Jeffries D, Allen S, Detection of severe protein-energy malnutrition by nurses in The Gambia, Arch Dis Child, 2004;89:181-184

35. Bairagi R, Effects of bias and random error in anthropometry and in age on estimation of malnutrition, Am J Epidemiol, 1986;123(1):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(3):159-66

37. Davis LE, Epidemiology of famine in the Nigerian crisis: Rapid evaluation of malnutrition by heigh and arm circumference in large populations, Am J Clin Nutr, 1971;24:358-364

38. Lowenstein MS, Phillips JF, Evaluation of arm circumference measurement for determining nutritional status of children and its use in an acute epidemic of malnutrition, Am J Clin Nutr, 1973;26:226-233

39. Arnhold R, The QUAC stick: A field measure used by the Quaker Service team, J Trop Pediatr, 1969;15:243-247

40. Shakir A, Morley D, Measuring malnutrition, Lancet, 1974;20:758-759

41. Shakir A, Arm circumference in the surveillance of protein-calorie malnutrition in Baghdad, Am J Clin Nutr, 1975;28:661-665

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(3):636-645

43. English M, Esamai F, Wasunna A, Were F, Ogutu B, Wamae A, Snow RW, Peshu N, Assessment of paediatric inpatient care in first referral level hospitals in 13 districts of Kenya, Lancet, 2004;12;363(9425):1948-53

44. MSF, Nutrition Guidelines, MSF, Paris, 1995

45. Rivers JPW, The nutritional biology of famine, Biological Society Series, Oxford University Press, Oxford, 1988:57-106

46. Burgess HJ, Burgess AP, The arm circumference as a public health index of protein-calorie malnutrition of early childhood, The Journal of Tropical Pediatrics, Dec 1969: 189

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(1):11-18

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(10):715-26

49. Berkley J, Mwangi I, Griffiths K, Ahmed I, Mithwani S, English M, Newton C, Maitland K, Assessment of severe malnutrition amongst hospitalized children in Rural Kenya: Comparison or weight for height and mid upper arm circumference, JAMA, 2005;294(5):591-597

50. Berkley J, Newton C, Maitland K, Severe Malnutrition Assessment in Children in Rural Kenya, JAMA, 2005;294(20):2577

51. Waterlow JC, Classification and definition of protein-calorie malnutrition, Br Med J, 1972;3:566-569

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, VALID International, 2005

53. Sommer A, Nutritional anthropometry and mortality risk, Am J Clin Nutr, 1981;34:2591-2560

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(2):162-6

55. Sandiford P, Paulin FH, Use of mid-upper-arm circumference for nutritional screening of refugees, Lancet, 1995;29;345(8957):1120

56. Pelletier DL, Frongillo EA, Habicht JP, Epidemiologic evidence for a potentiating effect of malnutrition on child mortality, American Journal of Public Health, 1983;83:1130-1133

57. Briend A, Golden MHN, Grellety Y, Prudhon C, Hailey P, Use of mid-upper-arm-circumference for nutritional screening of refugees, Lancet, 1995;29:1119-1120