3. Pregnant and lactating women

3.1 Introduction

3.1.1 *Background*

Anthropometric evaluation of nutritional status during the reproductive cycle, particularly during pregnancy, is a widely used, low-technology procedure that may be expected to generate much valuable information, yet it has seldom been rigorously evaluated (*1, 2*). The biological mechanism that underlies the relationship between women's nutritional status and reproductive outcomes is not fully understood except in extreme situations (e.g. famine).

Unlike nutritional evaluation during other periods of life, which is concerned only with the individual(s) in whom measurements are made, measurements made during pregnancy and lactation are expected to reflect both the nutritional status of the woman and, indirectly, growth of the fetus and, later, the quantity and quality of breast milk.

At the clinic level, anthropometric measurements are routinely made on all pregnant women at the time of first contact with the health services and several times thereafter. Information obtained is also routinely incorporated into medical records. The impact of these activities, in terms of benefits for the health of the mother and the fetus or newborn, remains to be demonstrated by randomized controlled trials and has recently been challenged (*2*).

Measurements taken early in pregnancy should be used to evaluate the nutritional status of the woman and to predict how well she can cope with the physiological demands of pregnancy. Unfortunately, this objective is usually neglected, despite clear evidence that, in developing countries, pregnancy and lactation represent a major nutritional drain on the mother (*3*). Among well nourished women, moreover, excessive weight gain during pregnancy, followed by only a brief period of lactation, will be associated with postpartum overweight, increasing the risk of chronic diseases later in life. Measuring a woman's height provides a proxy indicator of childhood growth and skeletal pelvic structure and a good predictor of the risk of cephalopelvic disproportion and obstructed labour, which is a major cause of maternal death in developing countries.

Thus, anthropometric measurements made during the reproductive period should be designed to evaluate women's capacity to deal with the physiological stress of pregnancy, and to identify those women who would benefit most from nutritional interventions.

Perhaps the most widespread use of anthropometric measurements during pregnancy has been in evaluating the risk of fetal growth retardation and selecting women or populations for nutritional interventions aimed at improving fetal growth or prolonging gestation. This application has unfortunately not lived up to expectations (*4, 5*).

Intrauterine growth retardation (IUGR) of nutritional etiology can be a consequence of both low availability of nutrients from a malnourished mother and poor placental transfer of nutrients from a relatively well nourished mother. The latter problem, however, is unlikely to be detected by maternal anthropometry. The association between maternal nutritional status and gestational age at birth is unclear (*4*). Moreover, the two negative outcomes of principal interest, fetal malnutrition or intrauterine growth retardation and prematurity, are heterogeneous syndromes (*6, 7*), with maternal nutrition being only one of the causative factors. Improving maternal nutritional status has a significant impact on birth weight only under extreme conditions.

Important methodological issues should be considered when the use of anthropometry during pregnancy is evaluated. There is a strong correlation between preterm delivery and inadequate maternal weight gain, but the fetal contribution to total maternal weight gain cannot be separately determined during pregnancy. Estimation of gestational age, a fundamental issue when considering repeated anthropometric measurements, requires special facilities not usually provided in clinical settings (*8*), and a high percentage of women are uncertain of the date of their last menstrual period. Failure to give appropriate consideration to issues of this nature has obscured the interpretation of research findings on maternal anthropometry and its clinical application.

In this report, the use and interpretation at the individual and population level of single and serial anthropometric measurements during the reproductive cycle are discussed in the context of both adequate and severely limited health service resources. This approach, combined with a critical evaluation of the literature, is expected to contribute to the most effective, yet practical, use of these important clinical and public health tools.

Implementation of screening programmes and referral systems based on anthropometric measurements during pregnancy may be more feasible than improving the socioeconomic conditions of the population, but should never be considered as a substitute for such improvements.

3.1.2 *Methodology*

Several recent publications have discussed in detail methodological issues related to anthropometry during pregnancy (*9, 10*). In anthropometric terms, pregnancy is unique in two respects: the period of observation is relatively brief and anthropometric indices change rapidly.

Prepregnancy values of maternal weight, height, or skinfold thickness are only seldom available, although similar values of height will be obtained regardless of when measurements are made. Body weight measured no more than 2 months before conception is an acceptable approximation of prepregnancy weight. If this value is unavailable, a proxy for measured prepregnancy weight may be based on maternal recall or on a measure-

ment made during the first trimester of pregnancy (*9*). Most epidemiological studies consider that use of recalled prepregnancy weight introduces recall bias, and that use of body weight measured in early pregnancy introduces first trimester weight-gain bias. However, in a study of a group of adolescents in the USA, recalled prepregnancy weight correlated closely with measured weight (*11*). Moreover, recalled prepregnancy weights obtained by hospital staff and then by research project staff were in close agreement, with an intraclass correlation coefficient of 0.95 (95% confidence interval 0.94-0.96) (*1*). Prepregnancy weight can be used as an indicator of the need for maternal weight gain and as a predictor of fetal growth, and may contribute to understanding of the biological mechanism of the interaction between nutrition and reproduction.

Total weight gain during pregnancy, perhaps the most commonly used maternal anthropometric indicator, is determined by subtracting prepregnancy weight (or weight in early pregnancy) from the weight in late pregnancy (usually measured just before delivery). Unfortunately, the value of anthropometry in late pregnancy for predicting risk or selecting individuals or populations for interventions is limited; measurements are made after most of the fetal growth has been achieved and interventions to increase birth weight are less effective. It is of greater value for decisions on referral of patients to appropriate facilities for labour, delivery, and neonatal care, and – during lactation – for selecting individuals for interventions.

Additional considerations relevant to components of the weight gain indicator include:

- accuracy of gestational age calculations
- fetal contribution to total weight gain
- the use of postpartum net weight gain vs. late pregnancy weight gain minus fetal weight
- rate of weight gain.

Length of gestation is most commonly estimated from the date of the last normal menstrual period (LMP) as recalled by the woman at the time of her first prenatal visit. The accuracy of this method and its potential for misclassification of growth-retarded and preterm infants have been extensively discussed in the literature (*1, 8, 12*). However, the effect of errors in recall of LMP or in calculation of gestational age on the rate of weight gain between two prenatal visits appears to be minimal after the first trimester, i.e. from 14 weeks to term (*9*). In clinical practice in developed countries and in research settings, measurements taken by ultrasound techniques early in pregnancy (16-18 weeks) could improve the accuracy of gestational age estimations. Agreement (to within 2 weeks) between gestational age estimated by early ultrasound measures such as biparietal diameter and femur length or newborn physical evaluation and gestational age estimated by the date of last menstrual period has been used to select the study populations for epidemiological

studies of maternal nutrition (*8, 13*). Where ultrasound is unavailable or women receive no prenatal care until the latter half of pregnancy, symphysis-fundus (SF) height and recalled time of first fetal movements may complement LMP as means of estimating gestational age.

The contribution made by fetus and placenta to total maternal weight gain is almost 40%, and represents approximately 9% of the weight gain before 10 weeks, 23% from 10 to 20 weeks, 41% from 20 to 30 weeks, and 54% from 30 to 40 weeks (*14*). There is a positive association between total weight gain and fetal growth (or duration of gestation); however, since total weight gain reflects both fetal weight and maternal tissue gain, the weight of the fetus is included in both sides of the prediction equation.

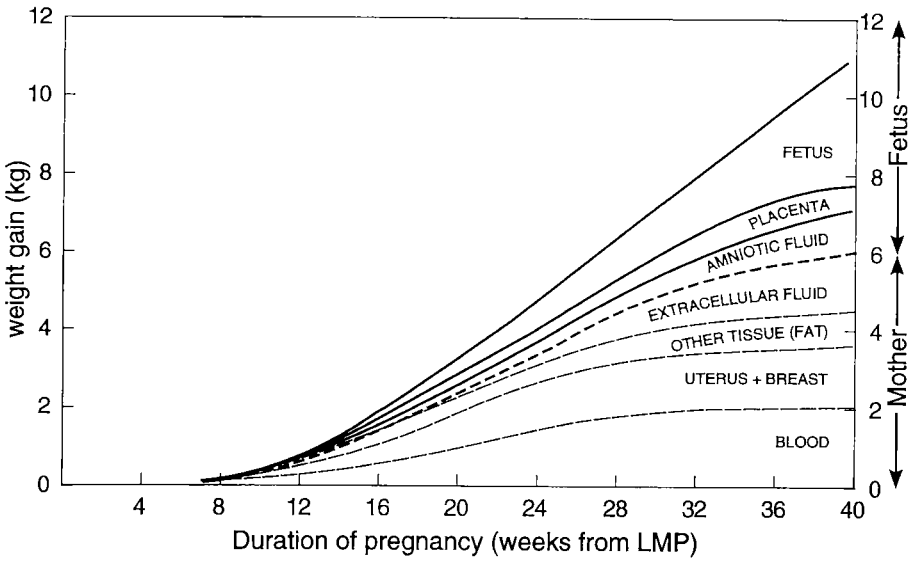
To eliminate the fetal contribution to total weight gain, the use of net weight gain has been suggested, obtained either by subtracting birth weight from total maternal weight gain (*9*) or by measuring maternal weight immediately after delivery (*6, 15*). The former method, however, takes no account of other products of conception or of maternal oedema, which together can represent up to 3 kg of the net weight gain (*6*). The second approach yields a measure of retained maternal weight. These approaches are of greater importance for research on the determinants of pregnancy outcome than for practical applications such as screening for intervention during pregnancy.

The extent of the correlation between total maternal weight gain and low birth weight (LBW) is probably distorted by the inclusion of preterm deliveries in many reports; the lower total weight gain in women who deliver preterm is likely to be a function of the shorter gestation. It has therefore been suggested that use of weight gain per week of gestation or rate of net weight gain during gestation has greater validity (*9*). Calculations can be made for the total gestational period, or more appropriately, by periods of pregnancy, if available. As is the case for total weight gain, the influence of fetal weight on the rate of weight gain will be less during the first part of pregnancy. Total or net weight gain is divided by the length of gestation and expressed as grams per week of gestation. During the period of linear weight deposition (from about 15 weeks to term) it is more appropriate, if tedious, to perform simple linear regression analysis, using three or more weight values for each woman, and to calculate the rate of weight gain from the linear regression coefficients of weight versus gestational age (*6*). Although this approach can be used for further statistical calculations in the context of research, it is impractical in clinical settings.

3.1.3 *Biological significance of anthropometry during pregnancy*

Some of the changes that occur during normal human pregnancy can alter the biological meaning of anthropometric measurements. Many of these changes relate to the growth of the fetus and of maternal tissue such as

Figure 8
Pattern and components of maternal weight gain during pregnancy[a]



[a] Source: Pitkin RM. Nutritional support in obstetrics and gynecology. *Clinical obstetrics and gynecology*, 1976, **19**:489-513. Reproduced with the permission of the publisher.

the breasts and uterus; others include the increases in body hydration and blood volume that occur quite early in pregnancy. The pattern and components of maternal weight gain during pregnancy are illustrated in Fig. 8.

3.1.4 *Anthropometry as an indicator of nutritional and health status*

Anthropometric indicators may be reflective of past events, predictive of future events, or indicative of current nutritional status. They may also indicate concurrent socioeconomic inequity, risk, or response to an intervention, or predict which individuals will benefit from an intervention. The distinctions between these different types of indicator are fundamental to their usefulness in the context of clinical application, programme implementation and management, and policy and planning.

Indicators of past and present status

Assessment of maternal status during pregnancy is commonly based on height, weight, mid-upper arm circumference, and various measures of skinfold thickness. In addition, maternal weight gain and fundal height may reflect fetal growth status.

Height in adults is a reflection of the interaction of genetic potential for growth and environmental factors that influence realization of that potential. In the more developed countries, genetic potential is the

primary determinant of height, since environmental constraints, such as acute and chronic disease, malnutrition, and socioeconomic deprivation, are minimized during the years of linear growth. In less developed countries, by contrast, much of the variation in adult height is the result of environmental influences on linear growth, especially those that affect growth in the first few years of life (*16*).

Use of maternal height as an indicator of health and nutritional status must therefore take account of the environmental context in which growth occurred. For example, a short woman in a developed country may be at risk of obstetric complications: her relatively small pelvis may be a constraint on vaginal delivery of a normally grown infant. A short woman in a less developed country, on the other hand, may be at high risk of bearing a poorly grown fetus if a poor childhood environment has persisted into her adult years, influencing her current pregnancy. The environmental conditions that lead to poor maternal linear growth may also result in poor growth and suboptimal development of the anatomical and physiological systems that sustain optimal fetal growth or maximize maternal health.

The biological changes that occur during pregnancy may affect the interpretation of maternal height relative to the non-pregnant state. The normal lordosis of pregnancy, for example, has been found to reduce maternal height as pregnancy progresses. This effect is significant enough to conceal increases in maternal height, associated with growth, in teenage mothers (*17*). In very young adolescents, in whom significant linear growth potential remains, some increase in height may be observed during pregnancy (*17*), but is likely to be very small (less than 1 cm). Adolescents may be misclassified as being at risk for poor pregnancy outcomes because of short stature relative to adults, when in fact the greater risk derives from other factors associated with adolescent pregnancy.

Body weight measured at various times during pregnancy has been widely used to assess maternal health status. In as much as weight is generally strongly correlated with height, it may serve as a general reflection of the past growth performance of the mother. However, since weight is changeable in adulthood and therefore variable within a given height category, this measure also reflects recent and concurrent health and nutritional status. Because body weight changes rapidly during pregnancy, gestational weight changes are routinely monitored as part of the prenatal care in many settings worldwide (*9, 10*). Interpretation of these weight changes is constrained by the fact that the various components of body weight may vary differentially depending on health and nutritional status, stage of gestation, and physiological condition, and according to genetic determinants. While total weight may be sensitive to these factors, it lacks specificity as an indicator.

The large variation in weight within a specific height category has given rise to various expressions of weight-for-height, such as the body mass

index (BMI). The exact significance of BMI is often difficult to determine. It may be used as an overweight index, on the assumption that excess weight for height reflects excess adiposity; however, while this may be valid for the upper extremes of BMI, it is less reliable for the middle of the population distribution in developed countries. In both developed and less developed countries, a very low BMI is a fairly accurate reflection of severe wasting of both fat and lean tissue (see sections 7 and 8).

Mid-upper arm circumference (MUAC) also reflects past and current status, but is less responsive than weight to short-term changes in health and nutritional conditions. It is relatively stable throughout pregnancy (*10*) and, even when measured relatively late in pregnancy, may be more reflective than weight of prepregnancy conditions. Other measures of limb circumference, such as of the calf (*18*) and thigh (*6*), have been proposed as indicators of status during pregnancy. These sites may involve more dynamic tissue, with changes in circumference reflecting changes in fat, muscle, and/or water specific to pregnancy. Oedema is increasingly common as pregnancy advances. Most pregnant women develop a degree of dependent oedema in the legs, which is considered normal during pregnancy (although pathological oedema can also occur); measurements of the lower body, specifically leg circumferences, may therefore be increased, particularly in late pregnancy.

Measurement of *skinfold thickness* at one or several sites is an increasingly common method of assessing nutritional status, but its use depends on several assumptions. First, skinfolds are assumed to reflect, at least to some degree, the overall distribution of subcutaneous fat. This approaches validity only if skinfold measurements are made at several sites. It is also assumed that the relationship between subcutaneous and total fat is sufficiently constant among populations (or that the factors that influence it are known and controllable) to allow total body fat to be estimated from skinfold measurements. In pregnancy, it is assumed that the relationship between skinfolds and total body fat described for non-pregnant women also applies, since normative data for pregnant women have not been reported. None of these assumptions, however, is likely to be universally valid.

In any individual, the proportion of body fat situated subcutaneously is variable according to certain influences (pregnant/non-pregnant, well/poorly nourished, etc.). In non-pregnant women with unusual fat distributions, skinfold measurements may thus yield very poor indications of total fat. In pregnancy, the situation is complicated by the influence of the various physiological changes on fat distribution and hence on skinfold thickness. Repeated measurements in pregnant women would therefore be unlikely to produce an accurate picture of the changes taking place in total body fat.

Relocation of existing fat stores from central to peripheral sites may occur during pregnancy to facilitate accommodation of the fetus in the

abdominal cavity. Thus, increased skinfold thickness on the arms, on the legs, or even on the back may not reflect an increase in the total body fat of a pregnant woman, although these sites are also believed to store the additional fat gained by many women during pregnancy. As pregnancy progresses, the enlarging abdomen makes it increasingly difficult to measure the abdominal skinfold reliably. The skinfold may appear very "thin" where it stretches over the uterine compartment, but since the abdominal skin is having to cover an increasing volume it may actually include an increased amount of subcutaneous fat compared with the prepregnant state.

As mentioned above, oedema may also affect skinfold measurements, particularly those made on the lower extremities in late pregnancy, by changing the composition and compressibility of subcutaneous adipose tissue. In undernourished populations, however, in whom the normal increases in plasma volume may be inhibited by malnutrition, the normal oedema of pregnancy may not be apparent.

Variation in skinfold thickness and circumference of the lower limbs may reflect the hydration status of the mother in middle to late pregnancy. Under the influence of increasing estrogen levels, changes in the water-holding capacity of subcutaneous ground substance allow more water to be held by the skinfold without overt oedema. This may increase the resistance of the skinfold to compression, resulting in an increased skinfold thickness, even when subcutaneous fat has not increased. Decreases in skinfold measurements observable in the first few weeks postpartum may reflect a reversion of the tissues to the non-pregnant hydration level, rather than an acute decline in fat stores.

Symphysis-fundus (SF) height has long been used to gauge the size of the pregnant uterus. Though originally used to assess gestational age (*19*), SF height has been evaluated as an indicator of fetal growth on the premise that the height of the uterus reflects its overall size and that this in turn reflects the size of uterine contents, which are dominated by the fetus by the second half of gestation. SF height has also been used to assess fetal growth deviations at both extremes (small and large for gestational age). In clinical settings its use has relied on single and serial measurements from mid-gestation to term. A recent review of 12 studies (*20*) concluded that SF height was of variable value as a predictor of IUGR. However, large values of SF height in late pregnancy have been used successfully to predict complications of delivery and problems with the newborn (*21*).

Indicators of risk, benefit, and response

In addition to their ability to reflect past and current health and nutritional status, all the measures discussed above have been shown to be related, with differing degrees of association, to various pregnancy outcomes. The ability of measures such as prepregnancy BMI and gestational weight gain to predict risk of LBW and pregnancy complications has led to their widespread acceptance as clinical tools. Later in this section,

results are presented of an extensive analysis of the degree to which some of these measures predict obstetric and neonatal risks. First, however, it is important to discuss the distinction between pregnancy-specific indicators of risk, indicators of benefit, and indicators of response. This requires some consideration of the role of anthropometry in the causal chain that leads from maternal health and nutritional status to pregnancy outcome.

One view of these relationships is presented in Fig. 9, taken from a recent report by the Institute of Medicine (*9*). This figure was developed to focus attention on the causes and consequences of variation in gestational weight gain. However, it also demonstrates the role played by other anthropometric factors in relation to pregnancy outcomes and as potential confounders or modifiers of the effects of weight gain on these outcomes.

The use of terms such as "determinant" and "consequence" implies causality, and all the relationships shown in Fig. 9 between maternal weight gain and its determinants or consequences are in some way assumed to be causal. This is important because any clinical or public health intervention designed to affect a particular anthropometric indicator will be ineffective in improving outcome for the mother or fetus/child if the associations between that indicator and the outcome(s) are not causal. Notwithstanding, the anthropometric indicator may still be valuable if it helps to identify women who might benefit from the intervention. Short maternal height in a chronically undernourished population, for example, may help to identify those at risk of LBW although it will not be influenced by an intervention that improves dietary energy intake during pregnancy. An indicator such as low weight gain during pregnancy, on the other hand, may identify women whose fetuses would benefit from a dietary intervention, and the intervention in turn may also improve the mothers' weight gain.

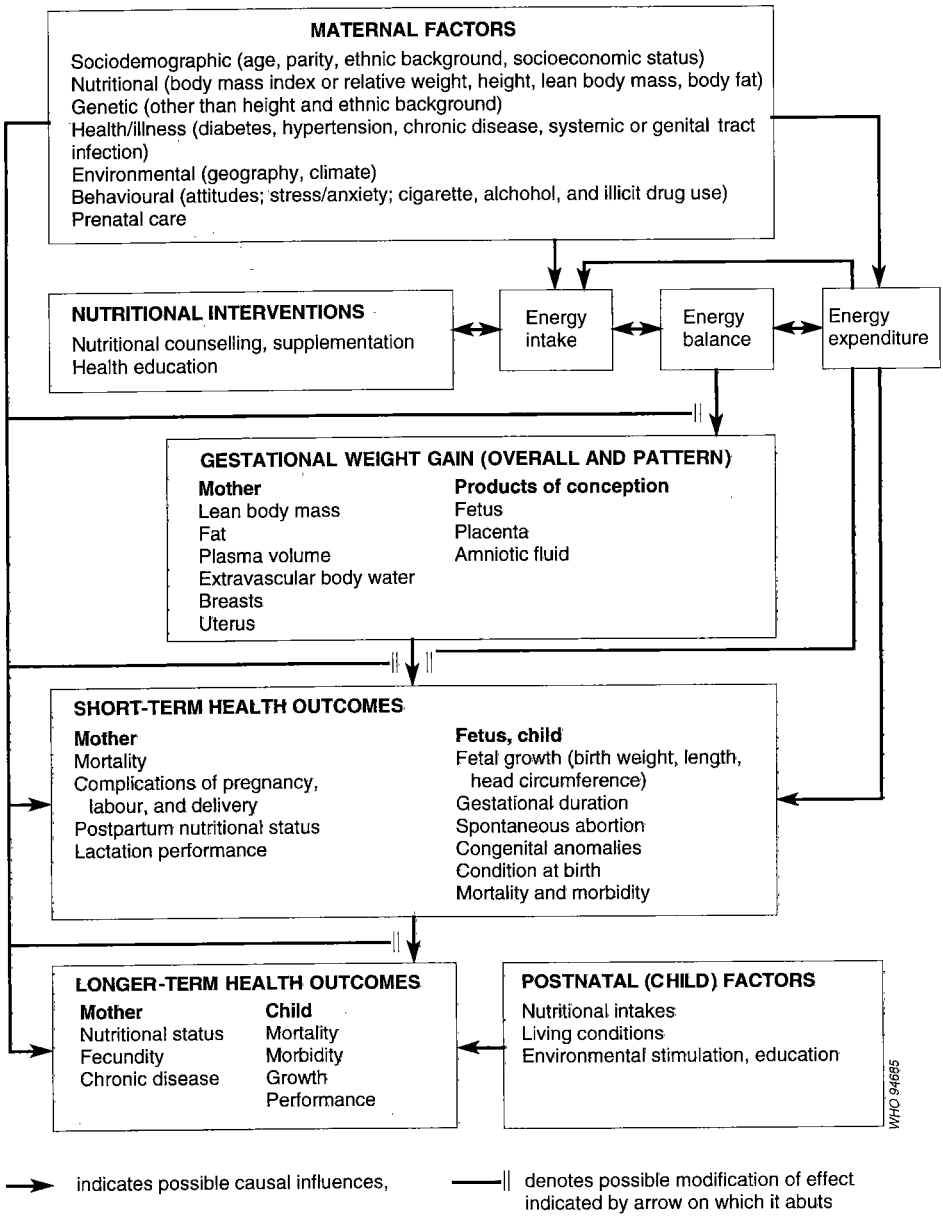
Another view of the distinctions between the concepts of risk, benefit, and response at the population level is illustrated in Fig. 10. As shown, maternal anthropometric deficits and physiological changes are among the several consequences of maternal malnutrition.

The more direct, or proximal, nutrition-related causes of these are past and current dietary intake, morbidity, physical activity, and the reproductive experience, which affect the state of maternal nutritional depletion or repletion. Non-nutritional causes include smoking and pregnancy-induced hypertension (PIH), which affect maternal physiology and neonatal development. More distal causes relate to maternal knowledge and behaviour, household resources, and environmental conditions. Although causal frameworks like this are commonplace, their more subtle implications for choosing and interpreting indicators are often overlooked.

Malnutrition as such is not shown in Fig. 10 because it is not directly measurable. Anthropometric deficits are often used as indicators of malnutrition, but for heuristic purposes it is best to consider these as consequences of poor diet and health status, along much the same lines as

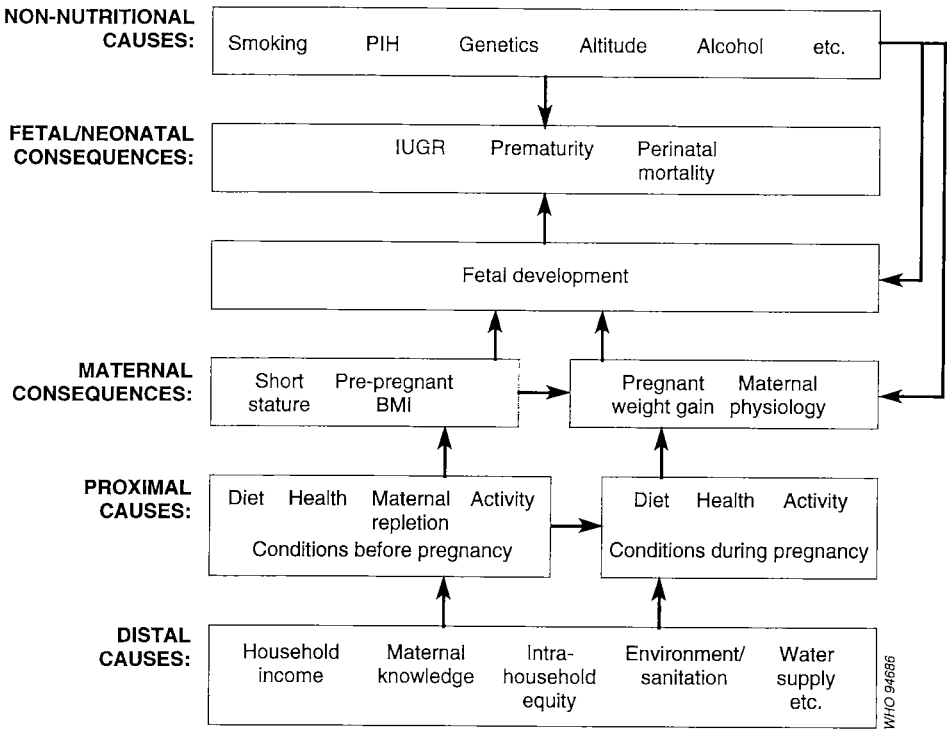
Figure 9

Schematic summary of potential determinants, consequences, and effect modifiers for maternal weight gain[a]



[a] Reproduced from reference 9 with permission from *Nutrition during pregnancy.* Copyright 1990 by the National Academy of Sciences. Courtesy of the National Academy Press, Washington, DC.

Figure 10
Selected causes and consequences of maternal malnutrition



other consequences. The common practice of placing malnutrition in the centre of such a diagram can be misleading when, as is often the case, malnutrition is assumed to be fully or largely captured in anthropometric characteristics. Conversely, the figure explicitly shows "maternal physiology" and "fetal development" as outcomes of nutritional and non-nutritional influences, even though they are not directly observable; they are important factors which may mediate, modify, or confound the relationships between causal indicators and indicators of consequences.

3.2 Using anthropometry in individuals

In choosing an indicator of maternal status during pregnancy from a number of candidate indicators, the crucial question is, "For what purpose will the indicator be used?"

3.2.1 *Choosing an indicator*

Anthropometric indicators have two primary uses – the targeting of interventions and the assessment of response to interventions. Matching indicators to application thus involves recognizing that a particular

disease is relevant, determining whether there is a treatment and how to identify (screen) women for the treatment, and evaluating the effectiveness of the treatment. The process of identifying the screening and response indicators necessarily involves assessing the degree to which a screening indicator identifies a subgroup that is more responsive to the treatment or, in epidemiological terms, the degree of effect modification achieved by the screening indicator. This can be done by several methods, but must recognize the prevalence of the poor outcome that is to be affected and the percentage of the population that is identified as at risk by the indicator. Finally, the selection of indicators will depend on practical and logistic considerations that include the availability of human and material resources, the timing and frequency of prenatal visits, and the acceptability of the measuring procedure to the women concerned.

This approach has been extensively developed for screening preschool-age children for malnutrition (*22*) and to a lesser degree for screening individual pregnant women (*10*). In a WHO collaborative study, this rationale also served as the basis for the analytical design applied to the evaluation of maternal anthropometry and pregnancy outcomes (*23*). This multicentre study supplements the findings in recent reviews on maternal anthropometry (*9, 10*) and on determinants of LBW (*24*), and serves as an excellent source of information to assist in the selection of appropriate maternal anthropometric indicators for screening.

None of the current literature has addressed the selection process as thoroughly and systematically as described above. Most of the previous emphasis has been on screening individuals at risk of poor pregnancy outcomes; the following discussion of indicators for screening individual women relies heavily on the published results of the WHO Collaborative Study with occasional reference to other supporting work.

The WHO Collaborative Study

The WHO Collaborative Study on Maternal Anthropometry and Pregnancy Outcomes (*23*) examined the relationships between maternal anthropometry and various pregnancy outcomes in separate studies of 25 population groups located throughout the world. Data were collected by local investigators from 1959 to 1989. They represent approximately 111 000 women for whom anthropometric data were collected repeatedly from early pregnancy to term. Among the various outcomes included in the data set, the most commonly reported are birth weight, gestational age, and pregnancy complications. The relative risks (or odds ratios where appropriate) of various deleterious outcomes based on maternal anthropometry were computed for each study. The results across studies were then combined into a meta-analysis.

The following outcomes were examined: LBW (<2500 g), preterm delivery (gestation < 37 weeks), small for gestational age (SGA) or IUGR (birth weight <10th percentile for gestational age), assisted (non-

spontaneous) delivery, pre-eclampsia (diastolic pressure >90 mmHg[1] with proteinuria and/or oedema), and postpartum haemorrhage (during first 24 hours). Fetal outcomes were fairly easily defined, but maternal complications were more difficult; in many studies there were no records of complications or the criteria (e.g. pre-eclampsia) were poorly specified. For these analyses, only the studies that yielded reliable data were included.

The maternal anthropometric indicators used were: height, weight, and BMI before pregnancy or during the first trimester, and at the 20th, 28th, and 36th weeks; weight gain between these various time points; and mid-upper arm circumference (MUAC) in early pregnancy.

Sample size for the 25 studies varied from 286 to 16 481, with a mean of about 4200; all but four studies had over 1000 subjects. Sample mean birth weights ranged from 2633 to 3355 g. The prevalence of LBW ranged from 4.2 to 28.2%, of SGA from 5.8 to 54.2%, and of preterm delivery from 4.6 to 56%. In the 14 studies that reported it, assisted delivery ranged from 2.2 to 27.6%; pre-eclampsia ranged from <1 to 15.4% in 11 studies, and postpartum haemorrhage from 0.5 to 4.4% in six studies. The individual studies produced similarly variable mean values for anthropometry. Mean maternal height ranged from 148 to 163 cm in 24 studies and mean prepregnancy weight from 42.1 to 65.6 kg in 23 studies. In 17 studies, mean weight at 20 weeks ranged from 44.0 to 64.6 kg, at 28 weeks from 45.5 to 65.3 kg, and at 36 weeks from 47.5 to 67.3 kg. Mean MUAC ranged from 21.9 to 25.4 cm in the 13 studies that reported it.

Odds ratios (ORs) were computed for each outcome by each indicator for each study that yielded relevant data. Since the values for each indicator varied greatly, the various studies were organized into between three and five groups, or clusters, identified according to similar statistical distributions of the indicator within the cluster. The distribution of the indicator was described for each cluster and the lowest 25% identified as the "at risk" portion of the sample.

Logistic regression was used to compute ORs for each study sample based on the risk of an outcome in the lowest quartile relative to the risk of the outcome in the highest three quartiles. The individual study ORs were then combined and an overall OR was computed for each of five pregnancy outcomes by each indicator.

The results of this analysis are summarized in Table 3. The first six indicators listed represent basic anthropometry in different stages of pregnancy. These are followed by a second set of derived indicators, such as BMI and weight gain, computed from the basic anthropometric data. The third set represents the results of an analysis of all the indicators in

[1] 90 mmHg = 12.0 kPa.

the first two sets, but only for those subjects in each study whose heights were below the population median height. This simulates a two-stage screening process, and is an indirect test of how the indicators would perform in the most stunted members of a population. Finally, the fourth set of results also simulates the two-stage screening, but uses prepregnancy weight below the population median weight as the first level of screening. The numbers enclosed in boxes represent combined ORs that are relatively high and for which the 95% confidence interval does not include 1.0, i.e. an indication of a significantly greater risk than expected by chance. Numbers that are underlined indicate a significant "protective effect" of the lower values of an indicator.

For almost all indicators, ORs are significantly greater than 1.0 for SGA and, therefore, for LBW. As pregnancy progresses towards term, ORs increase for measures of attained weight and weight gains from the prepregnancy period. The measures of weight gain appear to be more strongly related to risk of LBW and SGA if they include the initial weight measured at 20 weeks. Generally, the ORs increase substantially from the single-level screening to the two-level screening in which either short stature or low prepregnancy weight is considered as the first screening indicator. For predicting preterm delivery, measurements taken before pregnancy or in the first trimester have significant ORs in the range 1.20–1.49. Indicators related to weight gain vary widely, with some indicators (weight gain from 20 to 28 weeks) showing ORs between 1.43 and 1.86 and others (weight gain from prepregnancy to 20 or 28 weeks) showing ORs below 1.0, suggesting a protective effect of low cumulative weight gain.

Anthropometric indicators are less strongly related to complications of pregnancy and labour than to fetal growth. In general, ORs are less than 1.0; significant ORs suggest a reduced risk of complications in subjects with indicators in the lower quartile compared with the upper three quartiles of anthropometry. An important exception is the increased risk of assisted delivery in short women (OR = 1.61).

Interpretation of these results should take account of certain limitations to the analysis. Meta-analysis has been applied only recently in epidemiology, and there is still considerable debate about its validity and utility. In addition, potential bias in the combined ORs may be introduced by the method used to establish the cut-off points for each of the indicators examined in this study.

While ORs provide valuable quantitative estimates of the relationship between several anthropometric indicators and important outcomes of pregnancy, they represent only the first step in testing the utility of the indicator. The next step is to determine the optimum cut-off point for discriminating between women destined to have adverse outcomes and those destined to have favourable outcomes. This requires an analysis of the sensitivity, specificity, and positive predictive values (PPV) for those indicators with the best ORs (*25, 26*).

Table 3

Summary of estimated combined odds ratios (ORs) from the WHO Collaborative Study[a]

Note: Indicators with relatively high ORs for elevated risk are outlined; indicators predictive of low relative risk are underlined.

Predictors	IUGR or SGA	LBW	Preterm	Assisted delivery	Pre-eclampsia	Postpartum haemorrhage
Basic anthropometry						
Maternal height	1.91	1.72	1.20	1.61	0.88	0.72
Mid-upper arm circumference	1.63	1.93	1.22	0.88	0.69	0.65
Prepregnancy weight	2.55	2.38	1.42	1.00	0.71	0.71
Attained weight by week 20	2.77	2.43	0.99	1.04	—	0.96
Attained weight by week 28	3.03	2.41	0.89	0.91	0.87	0.97
Attained weight by week 36	3.09	2.59	—	0.87	0.71	0.68
Derived indicators						
Prepregnancy BMI	1.87	1.87	1.33	0.76	0.75	0.87
BMI by week 20	2.11	1.66	0.75	0.73	1.30	1.40
BMI by week 28	2.31	1.90	0.91	0.67	0.91	1.22
BMI by week 36	2.26	1.88	—	0.68	0.69	1.08
Weight gain: pp[b] to week 20	1.87	1.53	0.47	1.00	1.13	0.63
Weight gain: pp to week 28	1.85	1.53	0.78	0.75	0.82	0.81
Weight gain: pp to week 36	2.06	1.68	—	0.73	0.60	0.63
Weight gain: weeks 20 to 28	1.71	1.64	1.43	0.73	0.79	1.04
Weight gain: weeks 20 to 36	1.75	1.72	—	0.81	0.29	1.15
Weight gain: weeks 28 to 36	1.47	1.24	—	0.89	0.66	0.72

Table 3 *(continued)*

Predictors	IUGR or SGA	LBW	Preterm	Assisted delivery	Pre-eclampsia	Postpartum haemorrhage
Mothers of small stature						
Prepregnancy weight	2.99	2.63	1.49			
Attained weight by week 20	3.24	2.59	1.09			
Attained weight by week 28	3.56	2.65	0.95			
Attained weight by week 36	3.46	2.97	—			
Weight gain: pp[b] to week 20	2.79	1.96	—			
Weight gain: pp to week 28	2.85	2.09	—			
Weight gain: pp to week 36	3.20	2.30	—			
Weight gain: weeks 20 to 28	2.64	2.68	1.86			
Weight gain: weeks 20 to 36	2.67	2.82	—			
Weight gain: weeks 28 to 36	2.24	1.82	—			

Sensitivity, specificity, and PPV all depend on the relationship between a risk factor and a given outcome. However, the risk factor may be a statistical "marker" of the outcome without necessarily being the cause of it. To the extent that the success of an intervention depends on the causal link between the risk factor and the outcome, the etiological fraction (EF; also called the population attributable risk) will also be important. The EF is the proportion by which the incidence rate of the adverse outcome in a given population would be reduced if exposure to the risk factor were eliminated. It depends on both the relative risk (or OR) associated with exposure and the prevalence of exposure in the population. The EF is particularly important for those anthropometric indicators that will be targets for intervention, such as weight and BMI. For example, the EF associated with low gestational weight gain will indicate the maximum impact in reducing the incidence of a given outcome achievable by an intervention capable of ensuring adequate weight gain for all pregnant women in the population. EFs are therefore useful in defining the

Table 3 *(continued)*

Predictors	IUGR or SGA	LBW	Preterm	Assisted delivery	Pre-eclampsia	Postpartum haemorrhage
Mothers of low prepregnancy weight						
Attained weight by week 20	3.87	2.50	0.97			
Attained weight by week 28	4.02	2.75	1.07			
Attained weight by week 36	3.79	2.83	—			
Weight gain: pp[b] to week 20	5.58	2.71	—			
Weight gain: pp to week 28	5.36	3.49	—			
Weight gain: pp to week 36	5.63	3.36	—			
Weight gain: weeks 20 to 28	2.81	2.15	1.71			
Weight gain: weeks 20 to 36	2.49	1.68	—			
Weight gain: weeks 28 to 36	2.68	1.77	—			

[a] Reference *23*.
[b] pp = prepregnancy.

magnitude of expected effect of public health action designed to reduce adverse outcomes in a given community.

Finally, the feasibility of any risk assessment/management programme depends on the proportion of women who will be identified as at risk; adequate facilities, economic resources, and numbers of personnel must be available to deal with their subsequent referral and treatment. The anthropometric measurements chosen for risk assessment, as well as the cut-off points used to define risk, must take account of these practical aspects to avoid overburdening the health care system and to promote the efficient use of limited resources.

Relatively few studies have considered any of these factors, and none has considered all of them, in making recommendations for the use of maternal anthropometry. An analysis of misclassification was undertaken in the WHO Collaborative Study. While the analysis is rather restrictive and suffers from some of the same limitations as the estimates of ORs,

the results are informative. With cut-off values fixed at the 25th percentile of the cluster in which it was placed, each individual study was examined for sensitivity (SE) and specificity (SP) for each indicator relative to each outcome. If more than 40% of the studies met the criteria of SP>.7 and SE>.35 for a given indicator versus LBW and SGA (SE>.30 for preterm), the indicator was tagged as potentially useful for screening. The analysis is summarized in Table 4.

For predicting LBW, maternal prepregnancy weight and achieved weights at 20, 28, and 36 weeks performed equally well; about 50% of the studies met the criteria and had similar ORs in the range 2.4–2.6. When LBW is broken down into its components of SGA and preterm delivery, the results are generally as expected. The indicators that perform well in predicting LBW also perform well in predicting SGA, with similar sensitivity and slightly higher ORs. For predicting risk of preterm delivery, only prepregnancy weight and prepregnancy BMI met

Table 4
Summary of preliminary sensitivity (SE) and specificity (SP) analysis of various anthropometric indicators in pregnancy in 21 studies in the WHO Collaborative Study[a]

Anthropometric indicator	No. of studies[b]	% of studies meeting criteria[c]	Min. SE[d] (%)	Max. SE[e] (%)	OR (95% CI)[f]
Outcome = low birth weight					
Prepregnancy weight	21	62	35	47	2.38 (2.1-2.5)
Attained weight at 20 weeks	15	53	36	56	2.43 (2.0-2.8)
Attained weight at 28 weeks	14	50	38	52	2.41 (2.1-2.7)
Attained weight at 36 weeks	17	47	38	47	2.59 (2.2-2.9)
Outcome = small for gestational age					
Prepregnancy weight	20	50	36	52	2.55 (2.3-2.7)
Attained weight at 20 weeks	15	40	38	49	2.77 (2.3-3.2)
Attained weight at 36 weeks	17	53	36	50	3.09 (2.7-3.4)
Outcome = preterm delivery					
Prepregnancy weight	20	45	31	54	1.42 (1.3-1.5)
Prepregnancy BMI	20	40	31	39	1.33 (1.1-1.4)

[a] Reference *23*.
[b] Number of studies for which data exist for both outcome and indicator.
[c] Percentage of eligible studies with SP>.7 at cut-off of 25th percentile.
[d] Lowest sensitivity observed in studies with SP>.7.
[e] Highest sensitivity observed in studies with SP>.7.
[f] Combined odds ratio of all studies with 95% confidence interval (C I).

the criteria in over 40% of the studies. These same indicators had moderate combined ORs (1.33 and 1.42).

The major limitation of this analysis is that it does not allow for different cut-off values of the indicators across individual studies. The fact that the specificity criteria were met in 40% of the studies provides strong support for the general applicability of these predictors across many populations. However, the data should be re-examined using a series of common cut-off values, and the results compared across studies at the same cut-off values. This would help in determining the type of reference that should be developed: a reference with a common absolute cut-off, several references with a common relative cut-off value, or several (perhaps regional) references with different relative and absolute cut-off values.

Other evidence

The results of the analysis by the WHO Collaborative Study (*23*) generally confirm the findings of another meta-analysis in which the determinants of LBW were examined (*24*) (summarized in Table 5). This analysis relies on a thorough review of 895 studies published between 1970 and 1984. Strict criteria were applied to select studies with appropriate study design including variables that allow inference of potential causal determinants of LBW. However, relatively few studies provided the information needed to estimate relative risk (RR). Determined on the basis of low prepregnancy weight, the RR of preterm delivery was 1.25 compared with the OR of 1.42 in the WHO Collaborative Study. Kalkwarf (*27*) reports significant ORs of 1.42 (95% confidence interval = 1.25-1.60) and 1.37 (CI = 1.27-1.49) for prepregnancy BMI below 18.5 for 7312 white and 6730 black births, respectively, in the US National Collaborative Perinatal Project (NCPP). Kramer found no studies published before 1984 that showed a convincing relationship between the risk of preterm delivery and either maternal height or weight gain. However, in the NCPP Kalkwarf (*27*) has shown ORs of 1.65 (CI = 1.42-1.92) and 1.62 (CI = 1.46-1.79) for white and black women, respectively, for risk of preterm delivery when gestational weight gain was 100 g/week (the population 10th percentile) between 20 weeks and delivery. This compares with an OR of 1.43 (CI = 1.1-1.7) for weight gain between the 20th and 28th weeks in the WHO Collaborative Study.

In examining the risk of SGA from low anthropometric values, Kramer (*24*) reported RRs of 1.27 for maternal height, 1.84 for prepregnancy weight, and 1.98 for total gestational weight gain, compared with ORs from the WHO Collaborative Study of 1.91, 2.55, and 2.06, respectively. Kalkwarf (*27*) reported ORs of 1.83 and 1.44 for SGA in white and black women, respectively, from maternal prepregnancy BMI below 18.5, which are similar to the OR of 1.87 reported for the WHO Collaborative Study. Current evidence suggests the existence of biologically important correlations between many anthropometric indicators and fetal outcomes of pregnancy.

Table 5
Summary of anthropometric "determinants" of preterm delivery and intrauterine growth retardation[a]

Anthropometric indicator	Preterm delivery		IUGR[b]	
	RR[c] (no. of studies)	EF @ prevalence[d]	RR[c] (no. of studies)	EF @ prevalence[d]
Maternal height < 158 cm	1.0 (4)	–	1.27 (2)	6.3% @ .25 14.5% @ .63 18.5% @ .85
Prepregnancy weight < 54 kg	1.25 (3)	6.3% @ .27 10.3% @ .46 14.0% @ .65	1.85 (1)	11.9% @ .15 19.6% @ .29 28.7% @ .48
Total gestational weight gain < 7 kg	1.0 (1)	–	1.98 (2)	13.6% @ .16 36.6% @ .59

[a] Source: reference *24*.
[b] As inferred from SGA.
[c] Relative risk; number of studies meeting criteria for inclusion in meta-analysis is given in parentheses.
[d] Etiological fraction (%) at various prevalence rates; not computed if RR = 1.0, i.e. EF = 0.

In addition, Kramer (*24*) computed the EF for the three anthropometric indicators used in Table 5, reporting results only for those relationships found to be significant. Since EF varies as a function of prevalence of the risk factor in the population, results are reported for several prevalence rates. As an example of the interpretation of Table 5, intervention in a population where the prevalence of low body weight was 27%, resulting in the elimination of prepregnant weights below 54 kg, would reduce the incidence of preterm deliveries by 6.3%. Elimination of low prepregnant weight in a population of similar prevalence (.29) would result in a 19.6% reduction in the incidence of SGA.

3.2.2 *Applications of anthropometry for screening pregnant women*

On the basis of information provided by the WHO Collaborative Study (*23*), the detailed literature reviews of the Institute of Medicine report (*9*), and the Pan American Health Organization publication on maternal anthropometry (*10*), criteria were identified for the selection of nutritional indicators to be applied to individuals. These are listed in Table 6.

For anthropometric measurements made only once during pregnancy, generally at the time of a woman's first contact with the health care system, mid-upper arm circumference, height, weight (prepregnancy or early pregnancy, and attained weight at any stage during pregnancy), weight-for-height, and calf circumference were identified as of possible value in predicting maternal and fetal outcomes.

Table 6
Considerations in the selection of a nutritional indicator during pregnancy and lactation

1. Why is an indicator needed?
   - Screening for a nutritional intervention for the mother during pregnancy or lactation
     - to improve her nutritional state if she is underweight
     - to minimize complications during pregnancy (e.g. toxaemia, prolonged labour, need for assisted delivery)
     - to minimize maternal mortality.
   - Screening for a nutritional indicator to improve fetal health by:
     - reducing growth retardation
     - reducing preterm delivery
     - reducing morbidity/mortality.
2. Which instruments (scales, measuring tapes, etc.) are available?
3. What constraints are there on the availability of personnel and/or services?
4. Which nutritional indicators are available?
5. Is there any evidence of an association between these indicators and the outcomes of interest?
6. What are the biological bases for these associations? For what are these indicators proxy measures?
7. What is the minimum number of measurements required?
8. Are there reference data:
   - normative?
   - predictive of risk?
9. How are these references expressed? What data must be collected?
10. Is there any evidence that when applied to individuals or at clinic level the primary outcome is improved (randomized clinical trials)?

Mid-upper arm circumference is largely independent of gestational age and regarded as a proxy indicator of maternal prepregnancy weight or early pregnancy weight; it changes very little during pregnancy (*10*). Although the correlation between prepregnancy weight and MUAC is statistically significant, in most of the studies reported by WHO (*23*) this association is too weak to permit MUAC to substitute for prepregnancy BMI in individuals.

Table 7 lists sensitivity and specificity for several proposed MUAC cut-off points for the identification of pregnant women at risk of LBW, SGA, and neonatal morbidity. In Brazil, with an LBW rate of 23%, Lechtig (*28*) used a cut-off point of <23.5 cm and calculated a positive predictive value of 45% (n = 445, sensitivity 77%, specificity 71%). Thus, even where incidence of LBW is high, using this indicator to target interventions (to increase either birth weight or referrals for delivery at

Table 7
The use of maternal arm circumference for the prediction of neonatal outcomes[a]

Country	MUAC (cm)	RR	Sensitivity (%)	Specificity (%)	Outcome
Bangladesh	< 22.5	–	73	41	Neonatal morbidity
Brazil	< 23.5	–	77	71	Low birth weight
Chile	24	2.6	–	–	IUGR[b]
Guatemala	< 22.5 (14 days postpartum)	1.5	24	84	Low birth weight

[a] Adapted from reference *10* with permission. Copyright Pan American Health Organization, Washington, DC.
[b] As suggested by SGA.

tertiary centres) yields a high percentage of false-positive cases, i.e. women who are unlikely to benefit from the interventions.

Short maternal height has been associated with an increased risk of IUGR in several populations, and cut-off points between 140 cm and 150 cm have been proposed for screening. The WHO Collaborative Study (*23*) reports individual ORs and confidence intervals for the prediction of LBW from a cluster-specific definition of short maternal stature in each of 24 studies. Results are illustrated in Fig. 11. In most studies ORs are above 1.5 and the lower limits of the confidence interval are consistently above 1.0; exceptions are Guatemala, Lesotho, Malawi, and Viet Nam. The OR for all studies combined is 1.7 with a 95% confidence interval of 1.6–1.8 (see Table 3). Unfortunately, although specificity exceeded 70% in only four of the studies, sensitivity was very low (36–41%) for the detection of SGA at a height cut-off equal to the 25th percentile of a cluster-specific distribution.

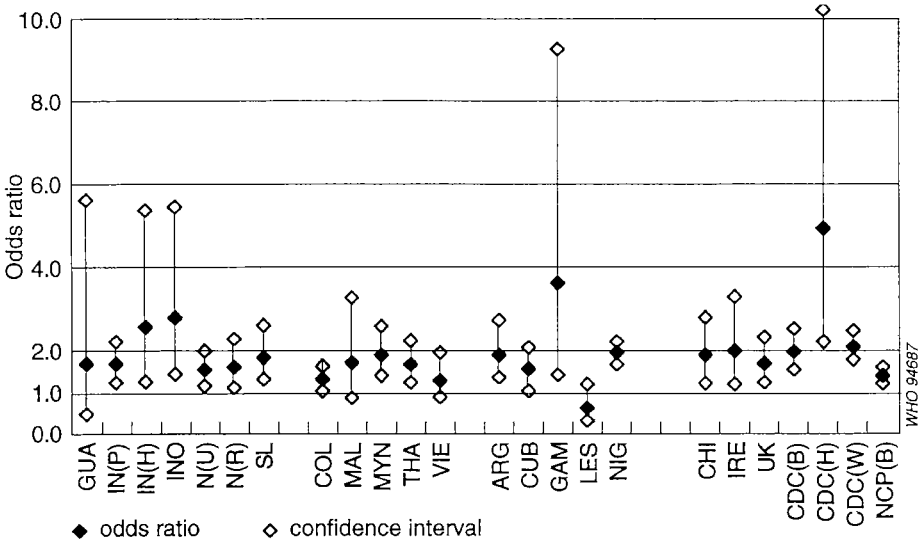
Short maternal height has also been shown to be associated with an increased risk of surgical delivery and intrapartum caesarean section among primigravidae (*29*). It could be of practical use for referring women for hospital delivery in areas where there are many home deliveries.

Results from the WHO Collaborative Study show that there is a strong correlation between a single measure of maternal weight late in pregnancy and SGA and LBW. While weight alone has a higher OR than BMI, it is a particularly good indicator when applied only to short women or those with low prepregnancy weights (Table 3). Rosso (*30*) also reports a strong relationship between weight-for-height at term and birth weight (Table 8). This indicator, even measured late in pregnancy, could be used for referral of women for delivery at facilities offering special care for the newborn.

Figure 11

Estimated odds ratios and 95% confidence intervals for low birth weight by maternal height[a]

Note: The ORs are for maternal height below the lowest quartile cut-off point vs. height above the highest quartile cut-off point.



[a] Source: reference *23*.

Key: GUA = Guatemala
IN(P) = India (Poona)
IN(H) = India (Hyderabad)
INO = Indonesia
N(U) = Nepal (urban)
N(R) = Nepal (rural)
SL = Sri Lanka
COL = Colombia
MAL = Malawi
MYN = Myanmar
THA = Thailand
VIE = Viet Nam
ARG = Argentina
CUB = Cuba
GAM = Gambia
LES = Lesotho
NIG = Nigeria
CHI = China
IRE = Ireland
UK = United Kingdom
CDC(B) = USA (blacks) (Centers for Disease Control)
CDC(H) = USA (Hispanics)
CDC(W) = USA (whites)
NCP(B) = USA (blacks) (National Collaborative Perinatal Project)

Much of the foregoing discussion of these studies is summarized in Table 9, which provides a structure for evaluating various anthropometric indicators for screening pregnant women. The table is divided into four parts, of which two (A and B) pertain to single examinations in settings with either low or adequate resources, the third (C) to multiple examinations where resources are adequate, and the fourth (D) to assessment of responses to interventions. Columns 1 and 2 deal with the purposes of screening or of evaluating an intervention and with what will actually be done for the individual woman as a result. The target populations and the type of anthropometric measurements are covered in columns 3 and 4, anthropometric indices and units of measurement in

Table 8
Influence on birth weight of weight-for-height of healthy, low-income Chilean women at term[a]

Weight-for-height (% of standard)	No. of subjects	Mean birth weight[b] (SD) (g)	Mean birth weight adjusted for height (g)
< 105	360	3224 (402)	3328
105-109	290	3219 (387)	3323
110-114	220	3264 (347)	3368
115-119	260	3327 (370)	3431
120-124	273	3419 (361)	3523
125-129	212	3501 (402)	3605
130-134	178	3486 (375)	3590
135-139	140	3555 (389)	3659
> 139	235	3642 (403)	3746

[a] Reproduced, with permission, from reference *10* (p. 179). Copyright Pan American Health Organization, Washington, DC.
[b] Liveborn, full-term infants; both sexes combined.

column 5, and the stage of pregnancy at which measurements should be made in column 6. Only the indices known to be related to the specific outcomes of interest are reported; in many cases, further evaluation of sensitivity, specificity, positive predictive value, and etiological fraction in specific applications and for particular situations would be necessary before final recommendations could be made. However, where current knowledge has identified a best index for a particular purpose, this is identified in the table.

Column 7 deals with reference data and criteria for judgement, i.e. the cut-offs applied to the various indices; these are not necessarily the references and cut-offs recommended by the Expert Committee, but those used by specific authors to generate the results summarized in column 8. Further information that may be relevant to the interpretation of anthropometric measurements is presented in column 9.

Where screening has involved multiple examinations of pregnant women (parts C and D of Table 9), the most common anthropometric measurement has been weight change. Depending on the purpose of screening, weight change may be determined over a short (several weeks) or long (the whole of gestation) period. Short-term weight change is generally determined early enough in pregnancy to allow for intervention to improve fetal growth; in the longer term, weight change may be used to identify women who should be referred to appropriate facilities for labour and delivery or for neonatal care.

In general, correlation between pregnancy outcome and weight change is weaker than that between outcome and a single determination of weight at any stage during pregnancy. This may be partly explained by greater error of measurement in determining weight change than in determining a single weight, and by the fact that a single weight determination also includes variation due to height. The WHO Collaborative Study (*23*) reported very high ORs (5.4–5.6) for the risk of SGA determined on the basis of weight gain from a low prepregnant value to 20, 28, or 36 weeks. Although many weight-gain charts have been used in prenatal care (*9, 31*), no single chart has proved to be useful over the whole range of applications covered by Table 9.

3.2.3 *Assessing response to an intervention*

Evaluation of the response of an individual woman to a nutritional or health intervention during pregnancy requires at least two anthropometric measurements – before and after the intervention. The indicator selected for the evaluation must be one that exhibits sufficient variation during pregnancy to be sensitive to intervention. Changes in values may be expressed as net difference between the two measurements, as percentage change from the baseline value, or as rate of change. Suitable indicators for the purpose are weight gain during pregnancy, changes in thigh and subscapular skinfolds, and changes in symphysis-fundus height, all of which are roughly linear after the first trimester of pregnancy (*6, 9, 32*). Normative data are available for these indicators, and the reports cited here present rates of change, with standard deviations, by week of gestation.

For clinical applications, it has been recommended that a sequential plot be made of maternal weight or weight increase for comparison with rate of weight gain of reference populations (*9*). Rate of weight gain can be calculated from the individual plots and compared with rates reported for normal healthy women who give birth to infants of normal birth weight (*6*).

Data on the effects of dietary changes on weight gain during pregnancy in six studies of maternal supplementation have recently been reviewed (*33*). It appears that unequivocal changes in maternal weight gain are seen only among women near starvation and, to a lesser degree, those who are clinically undernourished. For other women, even those at socioeconomic risk, the effects of these supplements on weight gain have been very modest and related to diet in early pregnancy, type of supplement, and prepregnancy weight (*33*). The use of these indicators for evaluating the impact of interventions aimed at improving maternal weight at the individual level require further evaluation, and it is not yet clear at this point whether the practice of weighing pregnant women at every prenatal visit should continue, particularly among healthy, well nourished women (*2*).

Table 9
Summary of recommendations for screening individual pregnant women for interventions

1	2	3	4	5
Uses: what will be done for the individual?	For what purpose?	Target group	What to measure	Indices and units
A. Low resource settings (no scales available), one examination to screen for poor pregnancy outcomes				
Refer for further evaluation if available; advise on or supplement diet	Prevent IUGR or treat newborn to prevent morbidity/death	Population at risk of IUGR	Mid-upper arm circumference	Absolute; cm (independent of gestational age)
			Calf circumference	Absolute; cm (independent of gestational age)
			Height[a]	Absolute; cm (independent of gestational age)
Refer for delivery at facility with neonatal care	Prevent neonatal morbidity/death due to prematurity	Population at risk of preterm delivery	Height[a]	Absolute; cm (independent of gestational age)
Refer for delivery	Prevent labour/delivery complications	Primiparas and history of dystocia	Height[a]	Absolute; cm (independent of gestational age)
Advise on or supplement mother's diet	Prevent depletion of maternal stores for lactation	Population with chronic undernutrition	Calf circumference	Absolute; cm (independent of gestional age)

6	7	8	9
Stage of pregnancy	Criteria for judgement (cut-offs)	Rationale for anthropometry	Other factors for interpretation
Any time during pregnancy	< 20.7 cm or < 23 cm; limited criteria for judgement (*23*, *10*)	For IUGR: RR=2.6 (*10*), RR=1.6 (*23*). Assesses body composition; proxy for prepregnancy weight	Limited value; modest sensitivity and specificity
Late in pregnancy	Not reported (*18*)	Best SE at 0.5 SP for IUGR (*18*); may reflect fat and lean tissue, also oedema (may be harmful or beneficial)	Promising, but not common measurement
Any time during pregnancy	Population-specific reference; likely cut-offs 147-153 cm (*23*)	RR=1.91 (*23*) RR=1.27(*24*) Proxy for past nutrition/ health status and SES	Limited use alone; poor sensitivity and specificity
Any time during pregnancy	Population-specific cut-offs,140-150 cm (*23*)	RR=1.00 for preterm (*24*) RR=1.20 (*23*)	Limited use; poor sensitivity and specificity
Any time during pregnancy	Population-specific cut-offs, 140-150 cm (*23*)	RR=1.62 for assisted delivery (*23*) Proxy for small pelvis	
Any time during pregnancy	None cited (*18*)	Identifies women who respond to postpartum supplement to improve lactation (*18*)	Mid-upper arm circumference may also be used but was not tested

Table 9 *(continued)*

1	2	3	4	5
Uses: what will be done for the individual?	For what purpose?	Target group	What to measure	Indices and units
B. Adequate resources (scales available), one examination (by early in second trimester) to screen for poor pregnancy outcomes				
Refer for further evaluation; advise on diet; dietary or medical intervention	Prevent IUGR or treat newborn to prevent morbidity/death[b]	Population at risk of IUGR	Weight (measured or recalled) and height	BMI
			Weight	Absolute; kg
			Weight	Weight for gestational age[a]; % of reference; kg
			Weight and height	BMI
Prevent preterm delivery or refer for neonatal care	Prevent newborn morbidity/death due to prematurity[b]	Population at risk of preterm delivery	Weight and height	BMI[a]
			Weight	Absolute; kg

6	7	8	9
Stage of pregnancy	Criteria for judgement (cut-offs)	Rationale for anthropometry	Other factors for interpretation
Measured during first trimester (or recalled early enough to intervene)	Population-specific 25th percentile (17-21 pre-pregnancy) (*23*)	RR=1.87 for BMI (*23*)	Recalled weights may not be reliable
	Population-specific 25th percentile (40-53 kg) (*23*)	For absolute weight: RR=2.55, high sensitivity and specificity (*23*); RR=1.84 (*24*). Indicates general body composition and health status	Absolute weight better than BMI; first screen for short stature improves RR for weight to 2.9 (*23*)
Measured during second trimester	Below % reference cut-offs (*9*), below population-specific 25th percentile (40-53 kg @ 20 weeks 43-57 kg @ 28 weeks) (*23*)	RR=2.77 @ 20 weeks, RR=3.03 @ 28 weeks, good sensitivity and specificity	First-stage screen for short stature improves RR for weight to 3.2-3.5 (*23*)
	Population-specific 25th percentile (19-22 @ 28 weeks)	RR=2.11-2.31 for BMI, good sensitivity and specificity (*23*)	Absolute weight better than BMI
Measured during first trimester (or recalled weight)	Population-specific 25th percentile (17-21 prepregnancy) (*23*)	RR=1.33, moderate sensitivity and specificity for prepregnant BMI	Best prediction in populations with lowest mean BMI
	Population-specific 25th percentile (40-53 prepregnancy) (*23*)	RR=1.42, moderate sensitivity and specificity for prepregnant weight (*23*); RR=1.25 (*24*)	Prediction not improved by first screen for short stature

Table 9 *(continued)*

1	2	3	4	5
Uses: what will be done for the individual?	For what purpose?	Target group	What to measure	Indices and units
Refer for delivery at facility with neonatal care	Prevent labour/delivery complications[b]	Population at risk of complications and of limited access to assisted delivery	Weight and height	BMI
Advise on or supplement diet	Prevent depletion of maternal stores for lactation[b]	Population with chronic undernutrition	Weight and height	BMI
C. Adequate resources (scales available), multiple examinations (by mid-pregnancy and at least twice before delivery) to screen for poor pregnancy outcomes				
Refer for further evaluation; advise on diet; dietary or medical intervention	Prevent IUGR[c]	Population at risk of IUGR	Weight	Absolute rate of weight gain; kg/week
				Absolute gain from prepregnancy weight; kg
			Weight and height	Absolute rate of weight gain; kg/week
	Prevent IUGR or refer for newborn care to prevent morbidity/death[c]	Population at risk of IUGR	Uterine (symphysis-fundus) height[d]	SF height relative to gestational age; cm

6	7	8	9
Stage of pregnancy	Criteria for judgement (cut-offs)	Rationale for anthropometry	Other factors for interpretation
Measured during first trimester (or recalled weight)	Population-specific 25th percentile (17-21 prepregnancy) (*23*)	RR=0.76 for BMI v. assisted delivery (*23*)	Height performs better than BMI. Note low BMI reported as "protective"
Any time during pregnancy	Below % reference population	Low body mass before or during pregnancy may persist after delivery	No observation found in literature
Up to 30 weeks' gestation	Population-specific 25th percentile (0.05-0.30 kg/ week between 20 and 28 weeks) (*23*)	RR=1.7 for weight change between 20 and 28 weeks	Not tested at other cut-off values
	Population-specific 25th percentile from weight gain chart (1 kg at 20 weeks, 3 kg at 28 weeks) (*23*)	RR=1.8 for total weight gain up to 20 or 28 weeks (*23*). Measures maternal and fetal tissue gain	
Up to 30 weeks' gestation	As above for weight gain between 20 and 28 weeks, and total weight gain to 20 or 28 weeks, but only in women of height or pre-pregnant weight below population mean	For short women: RR=2.64 for weight change from 20 to 28 weeks; RR=2.8 for total weight gain to 20 or 28 weeks (*23*). RR=5.6 for women with lowest prepregnant weight (*23*)	Not tested at other cut-off values
20-40 weeks' gestation	Below 10th percentile for gestational age (various reference data used) (*20*)	Proxy for fetal size. SE=0.5-0.7, SP=0.8-0.9 from 12 studies (*20*)	Useful in low resource area with multiple visits

Table 9 *(continued)*

1	2	3	4	5
Uses: what will be done for the individual?	For what purpose?	Target group	What to measure	Indices and units
Refer for further evaluation; advise on diet; dietary or medical intervention	Prevent preterm delivery[c]	Population at risk of preterm delivery	Weight[a]	Absolute rate of weight gain; kg/week
				Absolute total weight gain; kg
			Weight and height[a]	Absolute total weight gain; kg
Refer for delivery at facility with neonatal care	Prevent newborn morbidity/ death due to IUGR[e]	Population at risk of IUGR and with access to neonatal care	Weight[a]	Absolute total weight gain; kg
				Absolute rate of weight gain: kg/week

6	7	8	9
Stage of pregnancy	Criteria for judgement (cut-offs)	Rationale for anthropometry	Other factors for interpretation
20-40 weeks' gestation	Population-specific 25th percentile (0.05-0.3 kg/ week) (*23*)	RR=1.43 for weight change from 20 to 28 weeks, RR=0.47 for total weight gain to 20 weeks, poor sensitivity and specificity (*23*).	
	Population-specific 25th percentile (0-1 kg by 20 weeks) using provisional weight gain charts (*23*)	RR=1.6 for weight gain after 20 weeks (*27*).	Low total weight gain reported as "protective"
	As above, but in women below population mean height	RR=1.86 for weight change 20 to 28 weeks in shortest women (*23*)	
20-40 weeks' gestation	Population-specific 25th percentile (3.0-7.6 kg) for total weight gain to 36 weeks from provisional charts (*23*)	RR=2.06 for total pregnancy weight gain, RR=1.71 for weight change from 20 to 36 weeks	Allows use of measurements late in pregnancy, assuming there is a facility available for neonatal care
	Population-specific 25th percentile (0.05-0.32 kg/ week) for weight gain 20-36 weeks (*23*)	RR=3.20 for total pregnancy weight gain, RR=2.67 for rate of gain from 20 to 36 weeks in shortest women	
	First screen on height or prepregnant weight below population mean	RR=5.63 for total pregnancy weight gain, RR=2.49 for rate of gain from 20 to 36 weeks in women with lowest prepregnant weight (*23*)	

Table 9 *(continued)*

1	2	3	4	5
Uses: what will be done for the individual?	For what purpose?	Target group	What to measure	Indices and units
D. Assessing response to an intervention during pregnancy: adequate resources (scales available), mothers seen periodically from early pregnancy (examined at least twice before delivery)				
Refer for further evaluation; advise on diet; dietary or medical intervention	Prevent adverse effects on fetus (IUGR, preterm delivery)	Population at risk of adverse pregnancy outcomes	Weight, gestational age	Rate of weight gain[a]; kg/week
	Prevent maternal complications (e.g. toxaemia)	Population at risk of complications	Weight, gestational age	Rate of weight gain; kg/week
	Prevent depletion of maternal tissue for postpartum adaptations	Population at risk of under-nutrition or with low initial weight for gestational age	Weight, gestational age	Rate of weight gain; kg/week
			Thigh skinfold, gestational age (25–35 weeks)	Rate of fat gain; mm/week

[a] Indices/measurements recommended by the Expert Committee.
[b] Indicators listed in Part A may also be used.
[c] Indicators listed in Part B may also be used.
[d] Multiple measurements over time improve reliability.
[e] Indicators listed for preventing IUGR (column 2) may also be used.

Unfortunately, there are no published data documenting the rate of change in thigh and subscapular skinfolds or symphysis-fundus height after a nutritional or health intervention during pregnancy that would support its use for monitoring individuals.

3.3 Using anthropometry in populations

For many years anthropometric indicators have been used for assessing the nutritional status of populations in the context of surveys at national or community level, as a component of surveillance or monitoring

6	7	8	9
Stage of pregnancy	Criteria for judgement (cut-offs)	Rationale for anthropometry	Other factors for interpretation
20–35 weeks' gestation	Rate of weight gain greater than average (to allow for compensation)	Expected accumulation of maternal/fetal tissue due to intervention	Amount of weight gained will depend on the intervention and its effects on specific components (fat, muscle, water, fetus)
Mid-pregnancy to term	Rate of weight gain within normal range	Expected changes in hydration following treatment	
Throughout pregnancy (especially last trimester)	Sufficient excess weight gain to compensate for deficiency	Correction of inadequate balance of dietary intake, energy expenditure, and fetal demands, which leads to imbalance/depletion of fat and muscle	
	Positive gain		Subscapular skinfold may also be useful

systems, and for evaluating supplementary feeding programmes (*34*) or more general health/nutrition programmes (*35*). They are justifiably considered to be valid and practical indicators of the overall socioeconomic and environmental conditions of populations, especially young children, and have become increasingly accepted as such by international organizations and national governments (*22*).

While the widespread use of anthropometric indicators as tools in planning and policy-making is a positive trend in general, their value in programmes designed to improve the nutritional status of a population is

potentially far greater than is currently realized. Two major factors constrain achievement of this potential. First, there is a need for much greater perceptual clarity concerning the interpretation and use of anthropometric indicators for different purposes. Second, much of the research on anthropometric indicators fails to address the most urgent gaps in knowledge concerning their use in the context of policy and planning.

The conceptual issues discussed in this section relate primarily to maternal and fetal health, and research needs are highlighted as appropriate.

Figure 10 illustrates the relationship between nutritional and non-nutritional causes and consequences of maternal malnutrition and the way in which maternal anthropometric indicators may be used as measures of both outcomes and risk factors at the population level. Using a format identical to that of Table 9, Table 10 summarizes recommendations for the use of maternal anthropometry in populations. Much of the "evidence" that supports these uses is indirect and therefore extrapolated to the population level.

3.3.1 *Targeting interventions*

Targeting interventions to particular geographical areas or socioeconomic groups is the most common and best known application for anthropometric indicators, notably child indicators which are especially well suited to the purpose. Broad-based development programmes, for example, may well be targeted according to the prevalence of stunting among children, which closely reflects local socioeconomic conditions. In such cases, short stature is used as an indicator of socioeconomic inequity, and may often be combined with other considerations, such as literacy levels and housing quality (*36*). In economically disadvantaged populations, short stature in adults could also be used as an indicator of socioeconomic inequity.

As they relate to maternal anthropometry, socioeconomic indicators would include BMI among women (non-pregnant and non-lactating, or standardized for stage of pregnancy and lactation) as an overall indicator of the factors that affect women's energy balance (diet, workload, morbidity, reproductive demands). The importance of nutritional status as a factor in reproductive outcomes as well as maternal mortality makes a strong argument for the validity and usefulness of maternal BMI as an indicator of socioeconomic inequity. The same indicators may be used at the individual level to rank women according to degree of deprivation and to target resources to the most deprived, again using the underlying concept of socioeconomic inequity. However, in populations not characterized by energy deficiency (e.g. in developed countries), the significance of anthropometric indicators may be quite different; indeed,

the correlation between BMI and socioeconomic status of adult women in developed countries is likely to be the opposite of that in developing areas.

Indicators of socioeconomic inequity are the simplest to develop because they are required only to rank individual women or population groups from lowest to highest with respect to the measurement. The measurement may be chosen to reflect past inequities (height), recent inequities (weight-for-height in undernourished populations), or current inequities (dietary intake). The underlying assumption is that the measurement reflects some or all of the proximal or distal causes of maternal malnutrition (see Fig. 10), but no assumptions are made about the functional consequences of low indicator values or likely responses to the interventions proposed. For screening purposes, choice of cut-off points for inequity indicators may be governed strictly by the availability of resources for intervention.

By contrast, use of indicators of risk demands greater knowledge of the functional consequences of low indicator values. For instance, short maternal stature, low prepregnant BMI, and poor weight gain are all indicators of risk for IUGR, as are non-anthropometric factors like cigarette smoking and high altitude. Indicators of risk are often used when the principal concern is to prevent a particular adverse outcome (e.g. IUGR), or to ameliorate or prevent its consequences (e.g. neonatal morbidity or mortality related to IUGR). Risk indicators are preferred for the second of these purposes, unless the underlying reasons for the risk are well understood or can be ascertained, and are amenable to solution with the interventions available. They are often used for both purposes on the sometimes questionable assumption that causes are well known, can be ascertained, and are amenable to the available interventions.

Indicators of risk identify women who are more likely than average to have a specified outcome (e.g. IUGR); however, it does not necessarily follow that these women will benefit from the available interventions. For instance, short women are at risk for IUGR, but the degree to which they will benefit from supplementary feeding may actually depend less on height than on BMI. Risk indicators can be developed from observational studies, whereas predictors of benefit must be developed on the basis of intervention design and may vary according to the nature of the intervention.

These distinctions between indicators of inequity, risk, and benefit have important implications for how the indicators are used and interpreted. As mentioned above, inequity indicators can be used to target broad-based interventions designed to improve socioeconomic conditions. However, use of maternal height (as an indicator of risk for IUGR) to target more narrowly-based interventions (e.g. supplementary feeding programmes to prevent IUGR) may represent a misuse of the indicator:

Table 10

Summary of recommendations for screening populations of pregnant women for interventions and monitoring response[a]

1	2	3	4	5
Uses: what will be done for the individual?	For what purpose?	Target group	What to measure	Indices and parameters
A. Targeting of interventions: one examination				
Targeting for equity those deprived of access to social/ health/nutrition services	Ensure equitable access to services and reduce causes of maternal malnutrition	Population with socio-economic inequities and poor energy intake	Height	Absolute; means of functional groups[b]
			Weight, height, gestational age	Absolute weight or BMI; Z-score, % of gestation-specific mean[b]
Targeting interventions (supplementary feeding, vouchers, newborn care facilities, etc.) to those at risk of poor pregnancy outcomes	Prevent IUGR or provide access to neonatal care of IUGR infants; reduce presumed causes of maternal malnutrition and ameliorate poor outcomes for newborn or mother		MUAC	Absolute; prevalence below cut-off value (no correction for gestational age)
			Calf circumference	
			Height	
			Weight, height	Absolute weight or BMI
			Weight, height, gestational age	Absolute weight or BMI for gestational age[b]

6	7	8	9
Stage of pregnancy	Criteria for judgement (cut-offs)	Rationale for anthropometry	Other factors for interpretation
Any time during pregnancy	Rank by mean values for functional group/region; choice of intervention group depends on resources available	Measure of past inequity; socio-economic status related to anthro-pometry (*10*)	Cut-off to be age-adjusted for young teenagers
		Measure of current inequities (*9, 10*)	BMI preferred if recent inequities are to be considered
Any time during pregnancy	Cut-off may be population-specific, ranking by prevalence; choice of intervention groups depends on resources available	RR=2.6 for IUGR (*10*); assess body composition	Only as good as causal inference
		High sensitivity for IUGR (*18*); assess body composition	
		RR=1.9 (*23*); assess past nutrition/health and socioeconomic status	Cut-off to be age-adjusted for teenage mothers
First trimester measure or pre-pregnant recall		RR=1.84 for BMI RR=2.55 for weight (*23*)	Recalls unreliable; ab-solute weight includes effects of short stature
During second half of pregnancy		RR=2.77 @ 20 weeks, 3.03 @ 28 weeks for weight; RR=2.11 @ 20 weeks, 2.31 @ 28 weeks for BMI (*23*)	

Table 10 *(continued)*

1	2	3	4	5
Uses: what will be done for the individual?	For what purpose?	Target group	What to measure	Indices and parameters
Targeting interventions (supplementary feeding, vouchers, newborn care facilities, etc.) to maximize benefit by reducing poor pregnancy outcomes	Prevent IUGR or provide access to neonatal care of IUGR infants; reduce presumed causes of maternal malnutrition and ameliorate poor outcomes for newborn or mother	Population with socioeconomic inequities and poor energy intake	Weight, height, gestational age	Absolute weight or BMI for gestational age[b]
B. Assessing response to intervention: at least two examinations				
Evaluating long-term response to nutrition interventions, to allow subsequent modification	To ensure adequate nutrients are available to the fetus[c]; to add to mother's reserves for lactation	Women with nutritional deficits	Weight, gestational age	Weight gain, kg/week, over at least a 4-week period[b]; prevalence below cut-off; mean rate can also be used
			Thigh skinfold	Fat gain, mm/week, over at least a 4-week period[b]; prevalence below cut-off; mean rate can also be used

[a] The only uses listed in this table are for targeting interventions and assessing responses to interventions. Ascertaining determinants and consequences of malnutrition with the population as a unit of analysis has not been undertaken to any great extent. This type of ecological analysis is usually a first step to guide further studies at the individual level, which provide more definitive indication of the cause-effect relationship. Uses of nutritional surveillance are described in the text and are based on the same rationale as that presented in this table.

[b] Indices/measurements recommended by the Expert Committee.

[c] This assumes that, in the evaluation of most interventions, the response will be observed in the mother. The response is most likely to be reflected in the outcome of pregnancy; maternal anthropometric changes reflect intermediate or mediating mechanisms.

6	7	8	9
Stage of pregnancy	Criteria for judgement (cut-offs)	Rationale for anthropometry	Other factors for interpretation
	Cut-off may be population-specific, ranking by prevalence; choice of intervention groups depends on resources available	No data: possibly the same indicators as listed above for targeting risk	First stage screen for height improves RR for weight to 3.3-3.5 (*23*)
20-32 weeks' gestation or 2nd and 3rd trimesters	Change in means or prevalence of low values of weight gain	Significant maternal weight gain following supplementation in selected populations (*33*)	Total weight gain up to any date in late pregnancy is also useful if prepregnant weight is available
25-35 weeks' gestation	Change in means or prevalence of low values of weight gain	Rapid change in selected skinfolds during late pregnancy (*6*)	Subscapular skinfold may also be useful

maternal stunting reflects conditions that prevailed during the women's early childhood and may have little relevance to their current nutritional status. Thus, since maternal height is not necessarily predictive of benefit, supplementary feeding may not be an appropriate intervention. These theoretical considerations obviously need to be tested in appropriate research involving both maternal and neonatal outcomes (mortality and morbidity).

Targeting supplementary feeding programmes on the basis of maternal BMI involves three implicit assumptions:

- low maternal BMI is caused by chronically low energy intake by mothers in this population (rather than by morbidity-related factors shown in Fig. 10);
- low intake is caused by inadequate access to food at the household level (rather than by inequitable distribution of food within the household); and
- the food provided by the programme will be preferentially available to women (pregnant, lactating, or non-pregnant non-lactating) and will not substitute for the home diet.

An additional assumption is that supplementary feeding is preferred as an intervention to other measures that might be designed to reduce workload (e.g. labour-saving technology or changes in organization of labour in households and communities) or reproductive burden (child-spacing). Thus, the targeting of a particular intervention on the basis of anthropometric indicators presupposes that deficits in weight or height have specific, well understood causes that will be successfully dealt with by the intervention. In other words, assumptions are made about both causality and the efficacy of the chosen intervention.

Another aspect of the distinction between indicators of risk and benefit relates to the fact that two individuals (or populations) may attain the same value for a given measurement in different ways, which has strong implications for the types of intervention that would be most effective. For example, there is a well known relationship between birth weight and risk of mortality. Given that maternal diet and weight gain during pregnancy are determinants of birth weight, it might be expected that improved diet and weight gain would increase birth weight and decrease infant mortality correspondingly. However, it is also well known that the risk of neonatal mortality at a given birth weight is higher for preterm infants than for those who are SGA (*37, 38*). It is therefore unlikely that mortality in these two groups would be reduced to the same extent by a given intervention such as dietary supplementation. At a population level, it follows that the expected impact of improved maternal nutritional status on infant mortality will vary according to the pre-existing distribution of LBW across the SGA and preterm categories. This would explain, for instance, why dietary intervention during pregnancy would

have little or no impact on infant mortality in the USA (where most LBW is due to prematurity) but might have an important impact in populations with widespread and severe energy deficiency (where most LBW is a result of IUGR and a relatively large proportion of infant deaths can be attributed to SGA).

The above examples illustrate the general point that a given anthropometric indicator may be a valid indicator of inequity or risk, but will not necessarily predict benefit. Despite their profound implications, the distinctions between the different types are not generally appreciated by users of anthropometric indicators. As noted, development of indicators of benefit requires the use of intervention designs in which the differential impact of a given intervention can be examined across subgroups of women defined according to anthropometric indicators or other easily measured characteristics. Relatively little research has been done in this area, and apparently none that relates to maternal mortality as an outcome.

3.3.2 *Assessing response to an intervention*

Whereas predictors of benefit are useful for planning purposes in identifying the individuals or population groups who should receive a specific intervention, indicators of response are more valuable for assessing the effects of an intervention. An indicator useful for the one purpose may not necessarily be useful for the other. For instance, supplementary feeding for women of short stature (an indicator of risk for IUGR) will not increase their height but may well improve their weight gain; thus weight gain is the better indicator of response. A less obvious example is supplementary feeding of pregnant women with moderately low BMI (an indicator of risk); depending upon the degree of undernutrition in the population, the intervention may have more significant impact on birth weights than on any index of maternal anthropometry.

The concept of responsiveness of indicators is important because it suggests alternative ways of evaluating the impact of interventions on individuals and populations; it is also another factor to consider in interpreting the results of evaluation. For example, Beaton & Ghassemi (*34*) suggested that the failure of most studies to find any anthropometric impact of child feeding programmes might be explained by the extra energy provided by the dietary supplement being used for greater physical activity rather than for growth: though conjectural, this serves to underline the point that indicators used to screen individuals or target populations may not be those in which an impact is expressed.

Thus, anthropometric indicators need not lie on the causal pathway linking two events. They may simply be convenient markers of causal processes, which explains why they may indicate risk but may not be predictive of benefit or responsive to intervention. Low birth weight is a

good indicator of infant mortality risk, and in some populations a proportion of LBW is a result of maternal undernutrition; in certain settings, however, improvement of maternal nutritional status may have no significant impact on infant mortality (*39*). Similarly, weight gain may be poor among pregnant women in a population with a high incidence of morbidity during pregnancy, and the latter may have important effects on fetal development; in such a case, dietary intervention may well improve weight gain yet fail to improve the outcome of pregnancy.

3.3.3 *Ascertaining the determinants and consequences of malnutrition*

Anthropometric indicators are often used as outcome variables for analysing the determinants of malnutrition in research and planning settings. They are also used to ascertain the consequences of malnutrition; in many cases they are well suited to this purpose but, as in previous examples, problems may arise if anthropometry is equated too closely and uncritically with nutritional status itself.

Small-for-gestational-age, for example, has several non-nutritional causes (smoking, altitude, pre-eclampsia) and is also susceptible to various nutritional influences that operate at different stages in the mother's life. In any given country it may be relevant to health policy to determine the contribution made by maternal malnutrition to SGA. Short maternal stature, low prepregnant BMI, and poor weight gain during pregnancy all reflect maternal malnutrition, and the contribution of all three variables to the risk of SGA could be estimated by means of an observational study (implying the use of risk indicators). However, a study of the effects on fetal growth of dietary supplementation during pregnancy may seriously underestimate the contribution of maternal malnutrition to SGA, depending upon the relative influence of nutritional and non-nutritional factors in the local population and upon the effects on SGA of short stature and prepregnant BMI. It is also essential to consider the distribution of nutritional status in the population; dietary supplementation in a reasonably well nourished population would not be expected to have an effect on birth weight. Theoretically, therefore, a study of the contribution made by maternal malnutrition to SGA would require the normalization of nutritional status at all stages of the mother's development, from infancy (and probably *in utero*) through adulthood and pregnancy. This hypothetical study would inevitably yield different results from the observational study described above. The anthropometric indicators used in the latter case reflect a variety of non-nutritional socioeconomic and health problems with independent effects on SGA, and thus result in overestimation of the importance of nutrition. Risk indicators, on the other hand, do not necessarily reflect the direct effects that dietary intervention may have on fetal development, with no corresponding response in maternal anthropometry, and the importance of nutrition is therefore underestimated.

3.3.4 *Nutritional surveillance*

Most experience in the use of anthropometry for nutritional surveillance is based on children rather than on women in the reproductive years or on newborn infants. The one notable and important exception is low birth weight, which has been advocated as an important surveillance and general health indicator. This section relates the principles of inequity, risk, benefit, and response to the interpretation and use of LBW in surveillance systems, and to the potential uses of maternal anthropometry in different types of surveillance system.

It should be stressed that "surveillance" is used here primarily in the context of assisting decisions that affect populations rather than individuals; the use of anthropometry for patient screening and monitoring is covered in section 3.2.2. The focus is on three types of surveillance: for problem identification, for policy-making and planning, and for programme management and evaluation (*40*).

Birth weight

Because it can be interpreted as an indicator of inequity, risk, benefit, and response, low birth weight has numerous possible uses in surveillance. It reflects inequities in the conditions affecting women (throughout life, not only in pregnancy); it can predict (or be used as a proxy for) the risk of neonatal and infant mortality; it can predict which population groups may benefit from improved antenatal care of women and neonatal care of infants; and, assuming that interventions are well chosen and properly implemented, it can be a very responsive indicator for evaluation purposes. However, the potential also exists for misusing the indicator. This potential relates to the examples given earlier:

- LBW alone does not indicate the relative contribution made by prematurity and SGA;
- prematurity and SGA have different causes and will respond differently (or not at all) to various interventions; and
- prematurity and SGA have different consequences for the newborn and require different forms of neonatal and infant intervention.

So long as these distinctions are recognized, and the relative contributions of prematurity, SGA, and other, antecedent, causes are known, the indicator is valuable for targeting and evaluating interventions designed to prevent either LBW itself or its consequences (e.g. mortality). The value of LBW as an indicator can be further increased by careful selection of cut-off points and by means such as restricting its interpretation to term infants only.

Maternal anthropometry

Although maternal anthropometry has not been used for population surveillance to date, there are sound justifications for this application. There is growing recognition of the significance of maternal nutritional

status not only for successful reproduction but also for the health and social status of women in general. Acceptance of maternal anthropometric indicators as indicators also of socioeconomic inequity can strengthen the advocacy needed to translate this recognition into policy; implementation of policy, in turn, will be supported by the use of the indicators for targeting and evaluation of interventions. The primary application of maternal anthropometry is concerned mostly with conditions of undernutrition rather than overnutrition, although this depends on the context.

If interventions are to be targeted on the basis of a general concern about socioeconomic inequity, the mean of labile maternal measurements or prevalence of low values would appear useful. This assumes that maternal stature reflects past conditions too distant in time to relate reliably to current conditions. However, if interventions are to be targeted on the basis of risk of adverse outcomes, the choice of indicators, parameters (means vs. prevalence), and cut-off points becomes critical. In this case, low (or very low) maternal stature may well be a strong predictor of the risk of IUGR (or of delivery complications or maternal mortality) and may be preferred to more labile indicators. Maternal stature may also be a valuable indicator if the actions it is to guide relate to improving obstetric and neonatal care facilities to prevent or ameliorate the consequences of LBW (resulting from IUGR or preterm delivery). Stature may be less useful than maternal BMI or gestational weight gain if the actions relate to prevention of IUGR itself. For evaluating the impact of interventions, the best indicator is the one that is most responsive to the particular intervention; this would not be maternal stature (except perhaps in relation to early adolescent pregnancy), but could be one of several others, such as BMI at various stages in the reproductive cycle, gestational weight gain, or postpartum weight loss.

The principal distinction between the optimal maternal indicators of risk, benefit, or response may be determined more by the choice of cut-off point than by the choice of indicator *per se*. Depending on the situation, the cut-off point that is optimal for predicting risk (providing maximum sensitivity and specificity) may be higher or lower than that required to predict benefit. Although there has been insufficient research on this issue, the following considerations are obviously pivotal:

- the nature of the outcome (e.g. prematurity, IUGR, LBW, neonatal mortality, delivery complications, maternal mortality, postnatal maternal depletion);
- the nature of the intervention (e.g. energy supplementation, iron/folate supplementation, reduced workload, child-spacing, malaria prophylaxis);
- the distribution of the anthropometric measurement in the population (i.e. percentage of the population below various cut-off points);
- the prevalence of the outcome; and

- the importance in the population of the cause targeted by the intervention relative to other causes (i.e. the population attributable risk).

A single anthropometric indicator (e.g. weight at 20 weeks' gestation) may be used to predict who is at particular risk of a given adverse outcome (e.g. IUGR), predict who will benefit from a given intervention (e.g. energy supplementation), and identify those who have responded to the intervention (e.g. prolonged child-spacing since the previous birth), but the most efficient cut-off point for each will be different and will vary according to various combinations of outcomes and interventions. Unfortunately, in the absence of empirical evidence on this point, it is usually assumed that the same cut-off point is relevant for all purposes. For example, the cut-off point for LBW is generally set at 2500 g to predict infant mortality, but the same cut-off is often used uncritically to evaluate the impact of maternal dietary supplementation.

These considerations suggest that the impact of surveillance would be enhanced by greater knowledge of the efficiency of various indicators and cut-off points for predicting risk, benefit, and response. For example, a great deal of policy-making and planning is (or should be) prompted initially by concern about a particular pregnancy outcome (e.g. IUGR). To describe the distribution of risk for IUGR (or its correlate, SGA) in a population, a reasonable anthropometric indicator would be a maternal characteristic (and cut-off point) that most efficiently predicts the risk of IUGR (assuming that data on birth weight and/or gestational age themselves are not available). This information would be useful in allocating resources to programmes that ameliorate either the consequences of IUGR or the particular causes of IUGR in the population concerned. It would also be useful for allocating "block" resources to decentralized levels, leaving the decision concerning the most appropriate interventions to planners at those levels. In other circumstances, however, it may be desirable to allocate resources for a specific intervention (e.g. supplementary feeding). In this case the most appropriate indicator, rather than one of risk, would be a maternal characteristic (and cut-off point) that most efficiently predicts the populations most likely to benefit from supplementary feeding. Use of a combination of indicators, such as maternal height and weight in pregnancy, maternal prepregnancy BMI and weight gain, or maternal BMI and season (e.g. pre/post-harvest), may provide the best prediction in some situations. In evaluating the effects of specific programmes (e.g. dietary supplementation) over time, the characteristic (and cut-off point) should be one known to be responsive to the particular intervention.

Given that the maternal characteristics and cut-off points are reasonably similar for each of the applications discussed above, the "best" choice for surveillance purposes would be a single indicator, which would simplify the process. However, if there are large differences between the "most

efficient" characteristics and cut-off points for different purposes, the cost and complications of using several indicators may be fully justified by the improvements in targeting resources for maximum impact, although there has been insufficient research on this topic to guide a rational choice of indicators.

3.4 Population data management and analysis

3.4.1 *Sampling considerations*

At the population level the principal uses of anthropometry in the area of public health and social development are:

- to determine the nature and extent of nutrition-related problems;
- to target resources to population groups, on the basis of equity considerations according to the risk of abnormality, or according to the probability that the population would benefit from the available interventions; and
- to evaluate the response of the population to the chosen intervention(s).

It is important to distinguish these three applications of anthropometry from research on the development and testing of anthropometric indicators. Research applications may require samples that are representative of a known population group and longitudinal measurements on the same individuals over time. These requirements may or may not be relevant to public health applications involving anthropometric indicators.

3.4.2 *Problem identification*

Maternal anthropometric indicators are nonspecific and simply reflect, among other conditions, the nutritional status of women in the population. They do not, of themselves, indicate the causes of any imbalances, but may be used in conjunction with other information to confirm the existence of public health problems. Problems may include inadequate dietary intake of energy, protein, or other nutrients, and the excessive physical demands of morbidity, heavy workload, or high reproductive burdens. In formulating actions to improve maternal anthropometry it is important to understand the relative importance of these problems in the population. First, however, the emphasis should be on documenting the extent of each problem in order to promote awareness among policy-makers and stimulate interest in identifying the causes of and solutions to the problem.

For promoting awareness, samples can be cross-sectional and need not always be representative; longitudinal samples are necessary only if gestational weight gain is the indicator to be used, or if it is impossible for women to recall important events needed for data standardization. In documenting nutritional status during pregnancy, for example, it would

be important to know the gestational stage of each woman in the sample in order to compare the measurements with reference values for the same stage. If the conception date (or LMP) cannot be recalled with accuracy, it may be necessary to obtain longitudinal data, note the date of delivery, and use this as a reference point for assigning anthropometric measurements made earlier to the appropriate stages of pregnancy. Similar considerations apply to the postpartum and inter-partum periods (although to a lesser extent by 6–12 months postpartum).

The representativeness of samples may be an important issue in some settings where a high degree of bias is suspected in "convenience" samples (e.g. clinic populations), although this is not necessarily true when data are collected simply for the purpose of identifying problems.

In countries with high attendance at antenatal clinics, the absolute number of women with low anthropometric values may be sufficiently high to indicate the existence of a serious problem, even if there is sampling bias. Where clinic attendance is very low, it may be possible to determine the direction of bias among clinic attenders; for instance, comparison of maternal education levels may suggest that clinic attenders are generally among the more educated, and this will indicate the potential direction of bias in the prevalence of maternal malnutrition derived from clinic data. Data from referral hospitals may have extreme levels of bias (because patients are those who are very sick or at high risk) and might best be analysed separately from non-referral clinics.

3.4.3 *Policy and planning*

The principal decisions involved in policy-making and planning concern the targeting of resources towards population groups defined according to various criteria: physiological (pregnant, lactating, or neither), demographic (maternal age, parity, reproductive history), socioeconomic (occupation, ethnic group, income, social class), or geographical (region, district, etc.). Physiological and geographical criteria are the most commonly used. Each of these groups can be ranked according to the prevalence of maternal anthropometric deficits to determine which should receive priority "attention". Attention may take the form of specific interventions (e.g. supplementary feeding) if the causes of the deficits are well known, of a "package" of basic maternal interventions, or of more detailed investigation of the causes of deficits to facilitate selection and design of the most appropriate interventions.

Longitudinal samples are unnecessary for these purposes unless there is reason to suspect serious recall errors. Even where recall errors exist, the ranking of the various population groups may not be affected unless the degree of error varies across these groups. When the decisions to be made are based on prevalence trends over time, rather than on cross-sectional prevalence, time-series data will be necessary but need not rely on serial measurements in the same individuals.

The primary criterion for judging the importance of representativeness of samples is the degree to which it may affect decisions that are based on the ranking of population groups. If non-representative samples show the same direction and extent of bias in all geographical regions, targeting decisions will not be adversely affected. Similarly, if the direction and extent of bias do not change over time, targeting decisions based on time trends will not be adversely affected. Seasonal trends, however, represent a special case of lability over time: the composition of clinic samples may well vary according to seasonal incidence of morbidity and to constraints of travel time or accessibility that affect attendance at clinics. In practice, the direction and extent of bias, and its lability over time, are seldom known, and representative samples are consequently preferable.

In many practical settings it may actually be preferable to use clinic samples, despite the existence and variability of bias. This is because many of the resources to be targeted are intended to be delivered at the static health facilities rather than to the general population. In this case the relevant statistics concern the number of needy women who attend the clinics in each area and the types of service that they require. The number of needy women in the catchment area as a whole is largely irrelevant unless there are mobile clinics or other ways of reaching the non-attenders. A strong case can be made for using the available resources to reach the remote and underserved communities, but in reality few countries have the infrastructure necessary to support such an approach.

3.4.4 *Programme management and evaluation*

Anthropometric data have several potential uses in programme management and evaluation:

- monitoring the degree of coverage (percentage of needy being covered) and yield (percentage of those covered who are needy) achieved by the programme;
- monitoring the degree to which prevalence is changing and moving in the expected direction; and
- evaluating the net impact of the programme (i.e. changes in anthropometric indicators that are attributable to the programme).

The three applications have different sampling requirements and involve different design considerations.

Coverage and yield

Coverage and yield are common process indicators (*40*). Yield can be estimated from samples from the programme itself, whether or not these are representative. Coverage, by contrast, requires population-based (i.e. representative) samples from the catchment areas. By their nature, programme management decisions based on these statistics require certain longitudinal data, even if only in the form of a baseline assessment and periodic reassessments. Yield can be easily assessed on a regular basis, since it requires information only on programme

participants; coverage is probably assessed less frequently because of the need for community-based samples to achieve representativeness.

Adequacy evaluation

Adequacy evaluation can include process indicators like coverage and yield; it can also involve evaluation of "gross impact" – the overall change in outcome indicators of maternal anthropometry. Discussion here focuses on the latter aspect. Programme managers and administrators need to know whether the general trend in outcome indicators is in the expected direction; even if such a trend cannot be wholly attributed to the effects of the programme, it provides some indication of whether the programme is having the intended impact. The samples for analysis are drawn from among programme participants and are thus not necessarily representative of the population as a whole. Certain longitudinal data are essential. These may be obtained from several measurements of the same women at different stages in the programme cycle (on entry and at various points after entry), with the "delta value" (i.e. change in anthropometry) aggregated for all women of similar characteristics. They may also take the form of a single measurement on each woman at a particular point of interest in the programme cycle (e.g. on entry or at last visit). The specific requirements will vary according to the nature of the programme and the degree to which the design of the information system makes allowance for various sources of bias and confounding.

Impact evaluation

Impact evaluation differs from adequacy evaluation principally in attempting to determine the degree of change in outcome indicators that is directly attributable to the programme. It therefore involves more extensive sampling as well as a variety of other design and analytical considerations, although requirements will vary according to the desired level of plausibility in the results. Clearly, the cost of impact evaluation is closely related to the required level of plausibility.

Within this framework, the essential requirement of all impact evaluations is that they estimate the change in outcome indicators among programme participants compared with the levels of change among non-participants. Thus, non-participants must also be sampled, and measurements at two separate times are required in both groups to provide longitudinal data. In general, these samples should be chosen to be representative of a larger population of interest, although circumstances frequently do not allow this.

3.5 Methods of taking measurements

Measurements recommended for use during pregnancy and lactation are weight, height, mid-upper arm and calf circumferences, thigh skinfold thickness, and symphysis-fundus height. Techniques for taking these

measurements are described in Annex 2. For some of these measurements, pregnancy imposes certain constraints. Specific problems related to the reliability of skinfold thickness, for example, which is affected by problems related to compressibility, have already been discussed (see section 3.1.4). The measurement of SF height is unique to pregnancy and has been described by Belizán et al. (*32*), while measurement errors for weight, height, arm and calf circumference, and skinfold in non-pregnant women have been described elsewhere and generally apply to measurements made during pregnancy. Estimates of measurement error for most of the maternal anthropometric indicators have been reported by Villar et al. (*1*), with interclass correlations for height, weight, SF height, and gestational age in over 200 replicate measures of 0.82, 0.99, 0.92, and 0.98, respectively.

3.6 Sources and characteristics of reference data

The reference data needed to evaluate maternal anthropometric indicators are generally variable in quality; in some cases, there are no reference data. For certain indicators (height, attained weight, mid-upper arm and calf circumference) the interpretation is cross-sectional, while for others (weight gain, skinfold change) it is longitudinal through different stages of pregnancy. Nearly all reference data are normative, in that they are based either on data from the general population or, in the case of attained weight and weight gain, on data from women selected on the basis of favourable pregnancy outcomes. Only prepregnancy weight-for-height reference data have been established in populations with a proven health risk, and then only on the basis of long-term mortality (see sections 7 and 8). Few of the reference data have been based on pregnancy or postpartum outcomes, including maternal mortality, and none has analysed the differences in the distribution of indicators between mothers with favourable outcomes and those with unfavourable outcomes. Such an analysis is essential for the selection not only of appropriate normative data but also of the best cut-off values for each indicator relative to specific outcomes.

Existing reference data and proposals for the development of more useful references where necessary are discussed in the following section.

3.6.1 *Existing reference data*

Reference data for the assessment of anthropometry in pregnant women derive from several different sources, but there has been no attempt to standardize them.

Height

Height has generally been evaluated relative to local reference standards. In well nourished populations, most adults achieve their maximum linear growth potential by the age of about 18 years, and reference standards have been established from surveys of representative populations of

healthy, non-institutionalized individuals. In the USA, for example, the age- and sex-stratified reference data from the National Health and Nutrition Examination Surveys (NHANES) have been used (*41*). Since mean final achieved height varies across developed countries (*42*), locally derived reference data are probably necessary for some population groups. However, there has been no test of differences in the predictive value of specific height cut-off values for unfavourable pregnancy outcomes such as SGA, preterm delivery, or the need for assisted delivery across populations with similar or different mean heights.

Assessment of height in undernourished populations, of whom a large proportion fail to achieve maximum growth potential, poses different problems. Short stature in these settings is more likely to reflect past deficits, the causes of which still persist. Because variations in height between well nourished and undernourished populations have different causes, very different intervention strategies may be required to deal with unfavourable outcomes (SGA, preterm delivery) that are apparently similar yet perhaps of very different etiology. Moreover, the much higher prevalence of the risk factor (i.e. short height) in undernourished populations has implications for allocation of scarce resources in settings where many women are identified as at-risk. The best cut-off points may therefore differ according to outcome and underlying causes, as well as with the availability of interventions that focus on maternal height. This last factor is particularly important for population level assessments; the best cut-off for height will be the value that selects only the number of women for whom resources are sufficient to implement intervention. Of course, this assumes that height is in fact the best indicator of risk for the outcome of interest. In fact, for all the fetal outcomes examined by the WHO Collaborative Study (*23*), weight was generally a better indicator of risk than height. Only assisted delivery was predicted better by height than by weight. However, the WHO analysis suggests that initial screening on the basis of height can improve the predictive power of other anthropometric indicators. This two-stage screening requires further testing, first for misclassification (especially false-positives whose taller stature would exclude otherwise high-risk women from further consideration), and then for optimal cut-off values for height that ensure the least misclassification and the best correlation of the second-stage indicator with the outcome.

Prepregnancy or early pregnancy weight and weight-for-height

Reference data for prepregnancy or early pregnancy weight are generally derived from the same types of survey as height references. It should be recognized that weight used alone may also represent variation in height; as a result its predictive power will be somewhat stronger than that of weight adjusted for height. From the literature it is clear that larger mothers have larger babies; what is less clear are the independent effects of weight and height, and how they work together, perhaps with height

modifying the effect of weight on various outcomes. The results of the WHO Collaborative Study (*23*) suggest that both low prepregnancy or early pregnancy BMI and short stature are significant independent risk factors for LBW, SGA, and preterm delivery; the ORs for both indicators are similar in magnitude but not as high as those for weight alone (Table 3). For clinical and public health applications the use of weight alone may be sufficient, but for research into the biological relationships, the independent, combined, and interactive effects of weight and height require clarification.

The Institute of Medicine (*9*) recommends the use of BMI for assessment of prepregnant nutritional status in well nourished populations. The reference standards proposed are based on the Metropolitan Life Insurance Company tables of desirable body weight-for-height (*43*). The percentage of desirable weight-for-height is converted to BMI and the cut-off values that distinguish underweight and overweight from normal weight correspond to approximately 90% and 120% of the Metropolitan Life reference. These desirable weights-for-height are based on a functional relationship between weight and life expectancy, but there is no indication that they also correspond to levels of obstetric risk for the mother or fetus, or of postpartum maternal health or lactation performance.

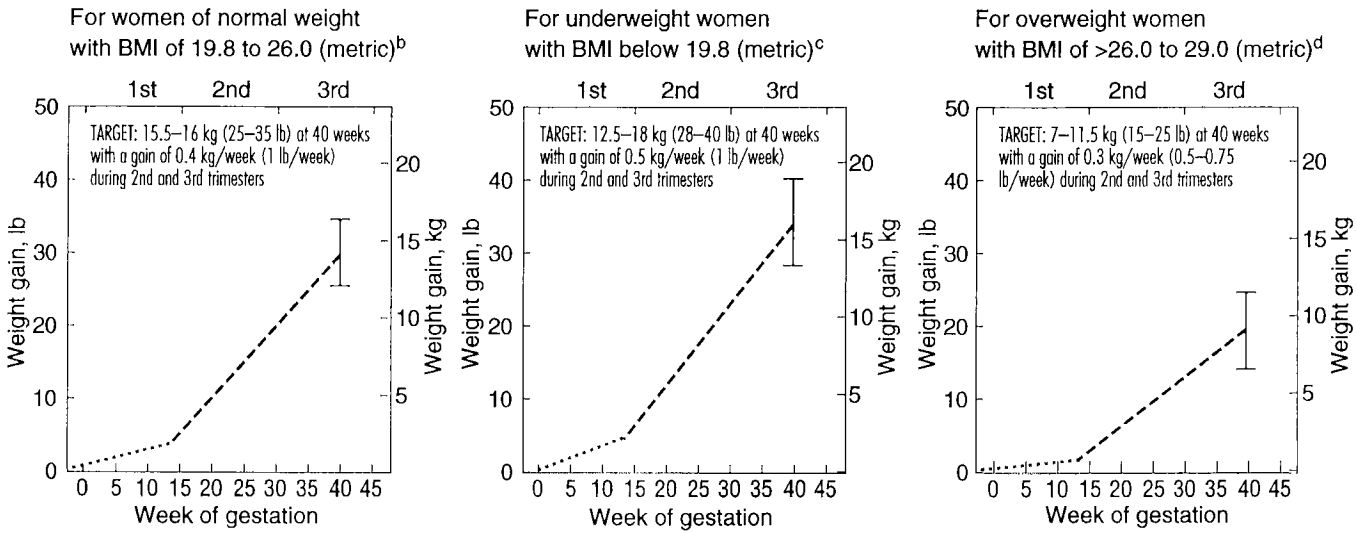
Attained weight during pregnancy

Attained weight measured at any time during pregnancy has generally been compared with reference weight gain charts. These charts have a long history, which has been described by Haas (*31*) and the Institute of Medicine (*9*). The first chart to be widely used was derived from data on 60 women from Philadelphia who were measured throughout pregnancy, gained an average of 10.9 kg by term, and had healthy pregnancy outcomes (*44*). A subsequent chart was based on 2868 healthy, primiparous British women with good pregnancy outcomes, who gained an average of 12.5 kg (*45*). These two became the basis for most weight gain charts that are used in the USA today.

The Institute of Medicine (*9*) has proposed a series of provisional charts (see Fig. 12) based on a nationally representative sample of US women who delivered full term (39–41 weeks), normally grown (3000–4000 g) infants without complications. The references were developed to reflect different weight gains associated with three categories of prepregnancy BMI. Ranges of accumulated weight gain are given only at term, since no data existed at the time on variation in achieved weights at various stages of pregnancy for an appropriate reference population. More recently, longitudinal data have been tabulated on 1185 women from San Francisco who had favourable pregnancy outcomes (Abrams, unpublished data reported in *31*). None of these normative reference charts has been tested for the overlap of distributions in women with favourable and unfavourable outcomes: the large variation in total weight

Figure 12

Provisional weight gain charts according to prepregnancy body mass index (BMI)[a]



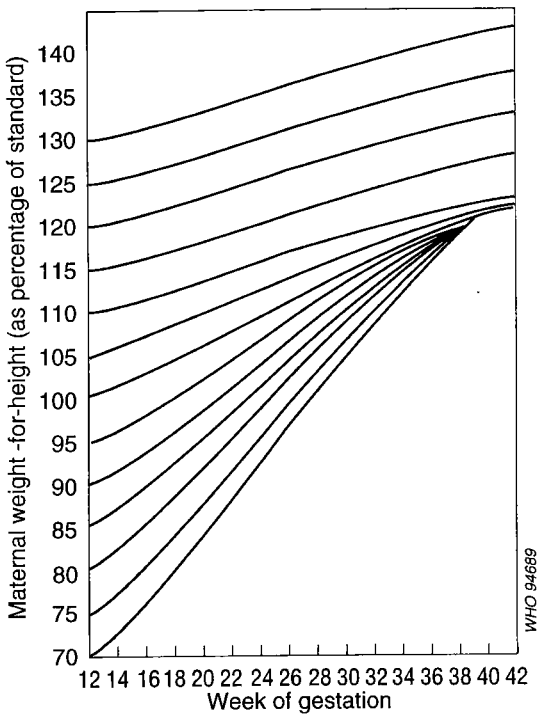
[a] Reproduced from reference 9 with permission from *Nutrition during pregnancy.* Copyright 1990 by the National Academy of Sciences. Courtesy of the National Academy Press, Washington, DC.

[b] Assumes a gain of 1.6 kg (3.5 lb) in the first trimester and the remaining gain at a rate of 0.44 kg/week (0.97 lb/week).

[c] Assumes a gain of 2.3 kg (5 lb) in the first trimester and the remaining gain at a rate of 0.49 kg/week (1.07 lb/week).

[d] Assumes a gain of 0.9 kg (2 lb) in the first trimester and the remaining gain at a rate of 0.3 kg/week (0.67 lb/week).

Figure 13
Chart designed for use in monitoring weight gain during pregnancy, considering prepregnancy weight-for-height as a percentage of a standard[a]



[a] Reproduced from reference *30* with the permission of the American Society for Clinical Nutrition.

gain from Institute of Medicine charts (*9*) suggests that it may not be possible to detect clearly different distributions in populations of this type. The 15th and 85th percentiles of total weight gain in the normative data used to produce the Institute of Medicine recommendations are 7.3 and 18.2 kg, respectively. In the San Francisco normative data the coefficient of variation in weight at various stages of pregnancy ranges from 90% at 13–14 weeks to 30% in the third trimester (Abrams, unpublished data reported in *31*).

A reference chart based on theoretical calculations of proportional weight gain has been proposed by Rosso (*30*) as an alternative to those based on normative data. This chart (see Fig. 13) is based on the assumption that total weight gain in most women should equal 20% of the ideal prepregnant weight-for-height. Use of the chart requires a knowledge of weight, height, and gestational age at any stage of pregnancy. Given height, ideal body weight can be estimated from a nomogram based on Metropolitan Life Insurance charts (*43*). Current body weight is then presented as a percentage of the ideal weight and

plotted on the curve (see Fig. 13). The application of these curves to undernourished populations, where prepregnancy BMI for a large proportion of women is below 18.5, is of questionable validity. For many of these women expectations of weight gain during pregnancy in order to compensate for prepregnancy deficits may be unrealistic. The curves are now being used to target women in Chile for a national food supplementation programme; evaluation results were not available at the time of preparation of this report.

The WHO Collaborative Study (*23*) provides insight into an approach for creating reference curves of attained weight. Though provisional, the curves produced show important features that are likely to be retained in the final version.

As described in section 3.2.1, the studies were first assigned to groups by cluster analysis according to the final weight achieved during pregnancy. Four clusters or country groups were identified, ranked from the lowest to the highest mean attained weight at 36 weeks of pregnancy. Each was then divided into three subgroups on the basis of birth weight (<2500 g, 2500–3000 g, and >3000 g). The weight gain curves for each country group according to birth weight are shown in Fig. 14. In all four, the curves for birth weights over 3000 g are clearly distinct from those for the two lighter birth-weight subgroups. Final attained maternal weights for the heaviest birth-weight subgroups are 55, 61, 65, and 73 kg for groups G1 to G4, respectively. It appears that women who deliver infants weighing over 3000 g have very different weights throughout pregnancy from mothers who deliver smaller infants, which suggests that this subgroup may be a suitable basis for constructing a normative reference.

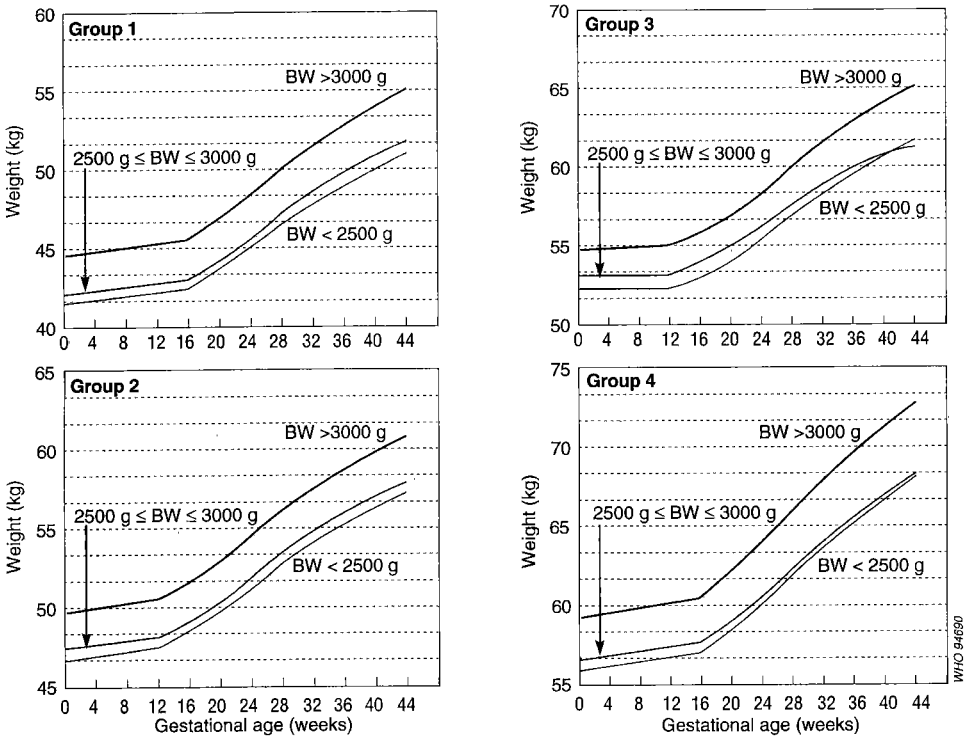
As a next step, assuming that the definition of the normative population for these indicators is satisfactory, more generic curves are prepared, to reflect accumulated weight gain from the beginning of pregnancy. The curves developed for each group are shown in Fig. 15; all four represent the weight gain pattern of women who delivered infants weighing over 3000 g. Of these, the three curves from groups of less developed countries (G1, G2, G3) are very similar, with a total weight gain of about 10.5 kg. In contrast, women from Ireland and the UK (G4) gain about 3 kg more, and their weight gain pattern is almost identical to that proposed as a reference by the Institute of Medicine (*9*). The similarity of the curves for groups G1, G2, and G3 might suggest that a single reference could be appropriate for these undernourished populations. However, it is important that this observation be validated with other outcomes (such as obstetric complications and lactation performance) before a definitive reference is recommended. Also, basing a reference on "positive deviants" (i.e. those who do well despite an environment where nutrition is not optimal, as in the countries of groups G1 to G3) is unwarranted (see section 2.9). It is also important to recognize any additional benefits of the extra 3 kg gained by the women in G4 (such as extra fat to

Figure 14

Pregnancy weight gain curves, by country group, for mothers with infants in different birth-weight (BW) categories[a]

Notes

1. The initial weight is based on the average prepregnant weight for that group
2. Group 1 = India (Poona), Nepal (rural), Nepal (urban), Sri Lanka
   Group 2 = Indonesia, Myanmar, Thailand, Viet Nam
   Group 3 = China, Colombia, Malawi
   Group 4 = Ireland, United Kingdom



[a] Source: reference *23*.

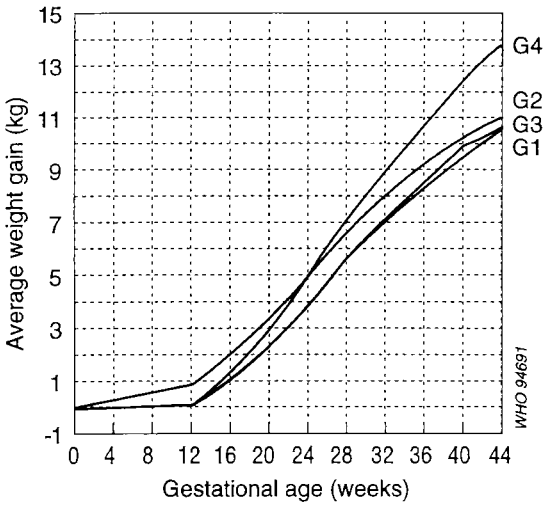
support lactation). Even though the pregnancy outcome is comparable across groups, G4 represents populations living in a "healthier" environment, which might lend support to this group serving as a normative reference. The final reference curve should present the median weight for gestational age of the normative population, along with lines representing -2, -1, +1, and +2 SD units.

It would then be desirable to test different cut-off values of maternal weight for their ability to identify women with poor pregnancy outcomes; sensitivity, specificity, and PPV would be examined at various cut-off values, at various gestational ages, relative to the risk of delivering an infant weighing less than 3000 g. Depending on the reasons for using maternal weight as the indicator (for screening for risk, response, or

Figure 15

Cumulative pregnancy weight gain by week of gestation for women delivering infants of birth weight >3000 g, by country group[a]

Note: G1, G2, G3, G4 refer to the country groups shown in Fig. 14.



[a] Source: reference 23.

benefit, or for estimating population characteristics), different cut-off points might be identified (see section 2) and could be added to the curves presented in Fig. 15.

In the meantime, a practical approach to using the results currently available is proposed, but should be fully field-tested and evaluated before being widely used.

Screening for risk of SGA to identify individuals for supplementary feeding can take advantage of the strong correlation shown in the WHO Collaborative Study between attained weight measured after mid-pregnancy and risk of SGA. As can be seen in Fig. 15, in countries with significant rates of undernutrition in women (i.e. groups G1, G2, and G3), the median weight gain was similar for all mothers whose infants' birth weights were above 3000 g: approximately 2.5, 6, and 8.5 kg from the time of conception to 20, 28, and 36 weeks, respectively. The 25th percentile was approximately 4 kg below the median in each group. For countries with no major undernutrition during pregnancy (group G4), median weight gain by mothers whose infants' birth weights were above 3000 g was 3, 7, and 10.5 kg, respectively, at 20, 28, and 36 weeks of gestation. The 25th percentile was about 6 kg below the median. Using these data to select the 25th percentile of attained weight at various stages of pregnancy, attained weight can be estimated from the following algorithm:

Estimation of 25th percentile attained weight at different stages of pregnancy derived from median weight of non-pregnant women (W_{np})

Level of undernutrition in pregnant women	Weight (kg) below which increased risk of SGA is expected		
	20 weeks	28 weeks	36 weeks
High	$W_{np} - 1.5$	$W_{np} + 2$	$W_{np} + 4.5$
Low	$W_{np} - 3$	$W_{np} + 1$	$W_{np} + 4.5$

This calculation requires the median weight, W_{np}, of non-pregnant women aged 20–29 years. This figure should be available (or readily measurable) for each ethnic group in each country. Use of the 25th percentile as a cut-off point is based on its association with an increased risk of SGA and its selection of a feasible proportion of women for treatment; in situations where more or fewer resources are available, the cut-off can be adjusted accordingly. Over time, a progessively smaller proportion of women should fall below this cut-off; initially it should be around 25%.

This approach to screening takes no account of other maternal characteristics, such as height, that may affect its efficiency. Results from the WHO Collaborative Study suggest that it should be possible to improve screening efficiency by first choosing women who fall below the population median height and then weighing them during pregnancy. These cut-off values should be subjected to an analysis of misclassification to test their efficiency in identifying women who would benefit from supplementation, and to see to what extent a two-stage selection procedure (height then weight) improves screening efficiency.

Weight gained during pregnancy

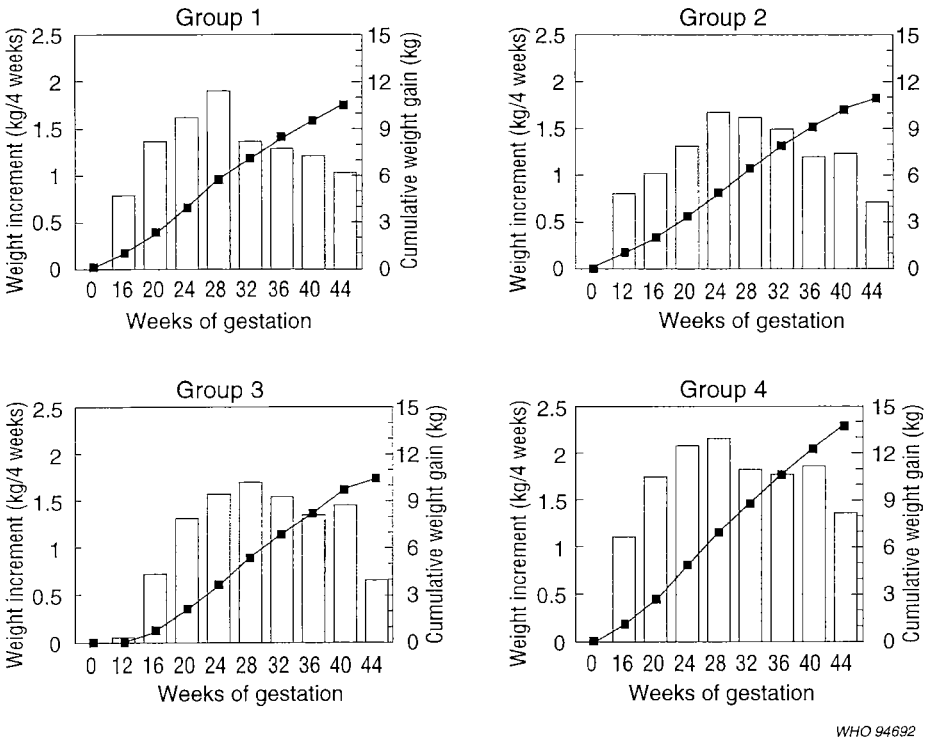
Although the charts described above are often referred to as "weight gain" charts they are actually "weight accumulated" charts in that they reflect what have been called "distance curves" (*46*). Actual weight gains should be expressed per unit time and plotted as velocity curves; only one true velocity curve has been published (*47*), and that relates to pregnant adolescents. A more common and practical expression of weight gain is the amount of weight gained between consecutive prenatal examinations and expressed as weight gain per week or per month. The Institute of Medicine (*9*) recommends a weekly gain of 0.4 kg during the second and third trimesters for women of normal prepregnant BMI, 0.5 kg for those who are underweight, and 0.3 kg for overweight women. Gains at the upper end of the range (3 kg/month) are suggested as reason for further evaluation of the mother; at the lower end of the range, gains of less than 0.5 kg/month by overweight women and less than 1 kg/month by women of normal weight should be cause for concern.

In the WHO Collaborative Study (*23*), preliminary analysis of monthly weight gain patterns in women delivering babies weighing more than

Figure 16

Mean monthly pregnancy weight gain and cumulative weight gain of mothers with infants of birth weight >3000 g[a]

Note: Groups 1, 2, 3, 4 are the country groups shown in Fig. 14.



[a] Source: reference *23*.

3000 g suggests consistency with the recommendations of the Institute of Medicine (Fig. 16). In developing countries gains of 1.5 kg/month during the last two trimesters are consistent with good pregnancy outcomes, while in developed countries gains of about 2.0 kg/month produce the same outcomes relative to adequate birth weight. Monthly weight gain seems to diminish somewhat in both populations from the second to the third trimester of pregnancy. No information on the distribution (standard deviation, percentiles) was reported for women with the more favourable outcomes, but means and standard deviations for all women, irrespective of pregnancy outcome, indicate considerable variation within clusters of populations (coefficient of variation 50 to 79% during mid-pregnancy). Villar et al. (*6*) report average gains of 375.1 g/week after 16 weeks of pregnancy for 105 healthy Guatemalan women who delivered infants with a mean birth weight of 3211 g (LBW = 4.8%), which is the same as that reported by the WHO Collaborative Study (*23*).

Mid-upper arm and calf circumference

Reference data for mid-upper arm circumference have been discussed at length by Krasovec & Anderson (*10*). Local reference data were generally used and based on cut-off values that range from 21 to 23 cm to identify women likely to have LBW infants. While MUAC seems to increase somewhat during pregnancy, the mean increases reported by Krasovec & Anderson are very small (generally less than 0.5 cm). These authors also suggest that MUAC measured at any time during pregnancy be used in place of maternal prepregnancy weight when scales are not available. This approach, however, should be validated before reference data are developed for application during pregnancy, since it implies that MUAC, like prepregnant BMI, may be used as both a screening indicator and a conditional factor for interpreting gestational weight gain. Recently, maternal calf circumference has been suggested as an effective screening indicator for risk of LBW and poor lactation performance (*18*), but no reference data specific to pregnancy are available and there is some question as to whether reference data for non-pregnant women are applicable. The lack of normative data from populations in developing countries where the measure may be most useful, and the fact that oedema is likely to affect the measure (improving its correlation with outcomes), suggests that more research is needed to evaluate the potential usefulness of this indicator. If possible, such research should include maternal mortality and morbidity outcomes in countries with poor health resources.

Skinfold thickness

Reference data on skinfold thickness specific to pregnancy have not been developed. In the relatively few studies that have used skinfolds throughout pregnancy (*3, 6, 9*), large changes in skinfold thickness at the medial anterior thigh site have been noted. Changes in skinfold thickness at many sites may be unrelated to body fat, especially during late pregnancy, and reference data based on non-pregnant relationships between skinfolds and total body fat are therefore not necessarily valid for extrapolation to body fat during pregnancy. By extension, the correlation between pregnancy outcomes and skinfold thickness (or its change during pregnancy) may not be simply a reflection of the mobilization of body fat stores to support fetal growth but may also involve other factors, such as changing hydration levels.

3.6.2 *Criteria for establishing reference data*

Criteria for describing normative reference data apply also to the establishment of functional cut-off points depending on local needs, resources, and applications.

Distance and velocity curves

Velocity is the rate of change of an anthropometric measurement and, by analogy, distance is the current value of that measurement (*46*). Although they are strongly correlated, the two should be considered separately.

Distance measurements are height, prepregnant weight, weight at any stage of pregnancy, MUAC, and skinfold thickness. Most weight gain charts are actually distance curves, indicating attained weight at a given point in pregnancy. However, if weight at conception is shown as zero, these curves can be considered as measuring velocity from time zero (usually date of LMP) to the point of measurement, provided that gain is divided by the elapsed time.

The most widely used velocity measurement is velocity of weight gain. For research applications, it is important to distinguish between gross gain and gain net of infant and products of conception; for simple prediction of risk, especially risk of LBW, IUGR, or macrosomia (rather than maternal nutritional status *per se*), the distinction is probably less relevant. Velocity of weight gain should also be adjusted for gestational age, since it is not linear throughout pregnancy. There is a wide choice of time intervals over which velocity can be measured, e.g. first, second, or third trimester, weekly, monthly, etc., which should be compared when charts are prepared. It is also important to consider measurement errors and timing of normal antenatal checks.

Velocity can be adjusted for, or considered independently of, distance; the two approaches may yield different results, so both should be investigated.

Conditional standard

An alternative to distance or velocity is a conditional or regression-based standard (*48*), which takes into account the possibility that optimal velocity is linearly related to distance, as proposed by Rosso (*30*). The conditional standard answers the question: What weight (or weight gain) is to be expected at the end of pregnancy, given the prepregnant weight?

Choice of reference data

Population. The ideal reference population is one in which the incidence of poor pregnancy outcomes is low. Assuming that obesity contributes to adverse pregnancy outcomes, this is unlikely to be a European or North American population, where overweight is a major problem. An African, Asian, or Central or South American population would be suitable, possibly one composed of relatively privileged people living in a healthy environment. Although the reference may be defined for a selected healthy population, it is essential that information is also collected on the population that is excluded by reason of poor pregnancy outcomes; this will permit the analysis necessary to establish functional cut-off points.

An alternative approach, and one that is easier to handle statistically, is to use two contrasting populations, one to define the lower limits of normal and the other the upper limits. The aim would be to identify ranges of anthropometric values within which birth outcome is generally good.

A further option would be a very large study, such as that of Naeye (*49*) with a sample size of 45 000, which used perinatal mortality rate as the

outcome measure. The large sample size ensures that the extremes of maternal height, weight-for-height, and weight gain are adequately represented. This is important if valid predictions are to be made for small and/or thin women. The data used were collected between 1959 and 1966 by the Collaborative Perinatal Project, and strongly influenced the 1970 recommendations of the US National Academy of Sciences on pregnancy weight gain. Although the data are now fairly old, they have the advantage of being unaffected by the recent trend to increasing obesity.

A further advantage is the use of perinatal mortality, rather than birth weight, as the outcome measure; since this is raised at both extremes of maternal anthropometry, the optimal central region can be identified unambiguously by fitting quadratic (U-shaped) regression curves relating mortality to anthropometry. The WHO Collaborative Study on maternal anthropometry (*23*) could also provide data to set these criteria, since it had both a large sample size (111 000) and a large number (25) of different populations from many different countries.

Study design. Ideally, reference data should be based on longitudinal studies, with anthropometric measurements made before and throughout pregnancy, and for 6-12 months postpartum. Measurements should be related to birth outcome, postnatal infant development, and maternal postpartum nutritional status. A cross-sectional study relating prepregnancy anthropometry to birth outcome would clarify the importance of achieved anthropometric values, particularly height, weight, and upper arm circumference.

Study size. For fully longitudinal studies with more common outcomes (SGA, preterm birth, etc.), a sample of about 1000 women may be adequate. A larger sample, of perhaps 2000, may be required for cross-sectional studies of similar outcomes. Perinatal and infant mortality studies require a very large sample, in excess of 10000. All of these sample sizes are dependent on the prevalence of the outcome in the population being studied.

Data. Desired measurements include height at the start of pregnancy (and at the end for adolescent populations), weight throughout, and arm and calf circumference, if possible. It is important to standardize measurement techniques, with periodic inter- and intra-observer comparisons. Data should also be collected on maternal race, parity, age, general health, and pregnancy complications, and on birth weight, gestational age, and sex of infants. Data of high quality are essential; this is especially true of gestational age, since any indicator that changes during pregnancy carries the potential for misclassification (*8*).

Analysis. Each available outcome measure should be used in turn for each of the analyses. Birth weight and gestational age should also be analysed as binary outcomes (LBW and preterm, respectively), and birth weight should be related to gestational age (small (SGA), appropriate (AGA), or large (LGA) for gestational age).

Cross-sectional regression analysis should be used to relate height and prepregnant weight to outcome. Distinctions should be made between weight alone, height alone, weight and height together, and weight corrected for height (i.e. weight-for-height). For the last of these, it is also possible to regress log weight on log height to determine whether BMI is the best weight-for-height index, or whether another power of height, e.g. $(\text{height})^3$, is better.

The results of this analysis can be used to identify ranges of weight and/or height associated with an acceptable outcome. It is important to look for non-linear relationships between anthropometry and outcome, particularly U-shaped curves related to mortality, to simplify the search for suitable cut-offs. For example, the probability of SGA or macrosomia as it relates to maternal weight and height can be modelled to identify the central region where the risks of both are low. Other maternal outcomes can be similarly combined to generate an optimal anthropometric profile.

The analyses should be extended to include velocity, particularly of weight gain during different periods of pregnancy. Velocity of gain in arm circumference, however, is unlikely to be informative.

Gross weight gain during pregnancy includes the weight of the fetus and products of conception. A more realistic impression of the correlation between weight gain and birth weight is obtained by using net rather than gross weight gain. However, this has little value in clinical or public health applications where prediction of adverse pregnancy outcomes is desired, since net gain cannot be calculated until *after* delivery.

The results of all analyses should be tested for their predictive power at the individual level by calculating sensitivity, specificity, and positive and negative predictive values. If these are not satisfactory there is little justification for putting forward anthropometric recommendations, unless it is made clear that they are for groups of women, not individuals.

3.6.3 *Recommendations for new reference data*

The recommendations in this section for reference data applied to specific indicators are based on the general principles discussed in the previous section.

Height

The reference for height of pregnant women should be cross-sectional and, in order to test for secular trends, especially in undernourished populations, and to examine when linear growth stops during adolescence, should sample across the reproductive age span of 15 to 50 years. The sample for the reference should represent healthy women who are likely to have reached their genetic growth potential, and who have had a favourable pregnancy outcome (gestational age between 37 and 42 weeks, birth weight between 3000 and 4000 g, no complications of pregnancy, labour, or delivery). It should be drawn from the general

population to which the reference is to be applied, so that height distribution can be compared for sub-populations with favourable outcomes (the specificity distribution) and sub-populations with specific unfavourable outcomes (the sensitivity distribution). This will allow the relationship between height and outcomes to be tested and the utility of height as an indicator of risk to be determined.

If height meets the criterion of significant correlation with outcome, the sensitivity, specificity and positive predictive value can be estimated from these data in order to compare height with other candidate indicators and to determine the best cut-off point for any particular proposed use (*25, 50*). This analysis should also take account of potential modification of the height/outcome relationship by maternal age, parity, and perhaps other known factors such as socioeconomic status or race that may serve as useful first-level screening criteria. For example, it may be necessary to use different cut-off values for adolescent and adult mothers in screening for risk of cephalopelvic disproportion or SGA, or height may prove to be a useful measure only for women of a specific population subgroup. The size of the sample drawn from the reference population must be large enough to allow the third or fifth percentiles to be established with confidence. Selection of the total sample size to test for cut-off points must take into consideration the number of cases of unfavourable outcome needed to produce clinically and statistically significant measures of association. Data that meet many of these requirements are available from the WHO Collaborative Study (*23*).

Height should be measured according to standard procedures (see Annex 2). Careful note should be made of major deviations in technique, such as whether the subject wore shoes, whether height was actually measured or simply recalled, and the certainty of age.

It would be useful to compare the results of this analysis with those from the general population of women of reproductive age in whom other short-term and long-term health risks are predicted from height.

Prepregnant weight or body mass index

Determination of appropriate reference standards for prepregnant BMI or weight should follow guidelines similar to those described for height. Since BMI reflects different etiologies and thus different associations for the same pregnancy outcomes, consideration should be given to the need for population-specific references. For example, in populations with marginal protein-energy nutritional status, variations in BMI reflect variations in lean body mass and all of its correlates (iron status, energy and protein reserves, physical activity, etc.). In contrast, BMI variation in populations with adequate protein and energy intakes generally reflects degrees of adiposity and obesity at one end of the distribution, and levels of lean body mass that are often related to greater physical fitness (less fatness) at the lower end of the distribution. The pregnancy-related consequences of variation in BMI may therefore differ widely between

populations. The contribution made by inter-population variation in body proportions (leg-to-trunk length ratio) to BMI variation, and the functional significance of this during pregnancy, is unresolved. Other issues that may have to be separately resolved for each population include the choice between weight and BMI as an appropriate indicator, which will also depend on the human resources needed to take and interpret the measurements.

Reference data for MUAC should follow similar guidelines.

Achieved weight and weight gain

Since serial measures are required for reference data on both achieved weight and weight gain during pregnancy, a longitudinal study is required that links achieved weight and/or weight increments with outcomes. Gestational age should be measured as accurately as possible: for application of the references, interpretation of achieved weight requires good estimates of gestational age. Interpretation of short-term weight gain does not require the same accuracy in gestational age, provided that increments are measured over short time periods (4–6 weeks). It is important to recognize that the relationship of weight or weight gain to outcomes may differ depending on specific characteristics of the mother, such as prepregnancy BMI, height, parity, age, and race. This should be formally tested to determine whether certain subgroups require separate references. Weight charts used to screen women at the first prenatal visit should be constructed with a series of mean and median weights at various weeks of gestation for a population with favourable pregnancy outcomes. To determine optimal cut-off values at each stage of pregnancy and for various outcomes, the trend line that is obtained by following these points should be bounded on both sides by lines of risk determined by the results of misclassification analysis. If optimal cut-off values have not been determined, -2, -1, +1, and +2 standard deviations (SD) should be plotted at each gestational age where a mean (or median) is known, and lines drawn to connect similar SDs across gestational age. These lines should be smoothed in ways similar to those described later for child growth curves (see section 5). This approach has been described by the Institute of Medicine (*9*).

Short-term weight gain may be expressed in two ways. A velocity curve can be used, similar to those used to assess child growth. The difficulty with this approach is that measures of variation or cut-off points of risk will vary with the length of the measurement interval; calculations then become necessary for every patient in whom measurements are taken over a different time frame from that used in the reference. Generally, where health care personnel have only minimal education and training, velocity curves are not well accepted. A more acceptable approach is to prepare a table of "optimal" weight gain that uses short (daily or weekly) intervals from which the health worker can easily calculate optimal gain over the time interval he or she is using. Alternatively, the table may be

designed with recommendations expressed as various options depending on the elapsed time between measurements; for example, there may be separate columns for observations made every 1, 2, 3, or 4 weeks. Regardless of how the recommendations are presented, however, it is essential that they include upper and lower limits of risk or SDs based on analysis similar to that described above for other indicators.

Skinfold thickness

Reference data for skinfold thickness can be constructed following similar guidelines to those described for achieved weight, and perhaps for weight gain, if a case can be made for a good association between skinfold change and specific outcomes. A major limitation of this approach is that it could result in a large number of reference curves or tables, each one specific to a certain sub-population and to certain outcomes. It would be valuable, but possibly difficult, to develop criteria that allow multipurpose reference data to be used with maximal efficiency.

3.7 Relationship between normative reference data and functional outcomes

The distinctions between indicators of socioeconomic inequity, risk, benefit, and response – discussed earlier – have profound implications for the construction of reference data for anthropometry at all stages in the life cycle, but little systematic attention has been given to these by the international bodies responsible for constructing reference data and advising on their appropriate uses. This section clarifies the issue and suggests the type of research required for further development.

The concept that underlies the construction and interpretation of current reference data is based on the specificity distribution for anthropometric characteristics. In other words, current reference data describe the distribution of anthropometric traits in an ostensibly healthy, well nourished population with favourable pregnancy outcomes. This concept has the greatest validity in the case of height of preschool children because it is possible to identify populations (e.g. the USA) in which the observed distribution and its moments can be assumed to reflect the variation in genetic potential within a population reasonably free of disease and environmental deprivation. In the past, this approach has also been used more uncritically for weight and weight-for-height of preschool children, although the existence of obesity among US infants and children is now recognized and invalidates the assumption that the population is "healthy and well nourished" in this respect. The definition of the specificity distribution for weight, weight-for-height, and related characteristics, in contrast to that for height, will require more knowledge about the long-term consequences of deviations in these characteristics at early ages. The situation is similar for anthropometry during pregnancy where short-term (pregnancy outcomes, maternal morbidity and

mortality) and long-term (maternal depletion, obesity later in life) consequences need to be examined.

It should be emphasized that the specificity distribution indicates only the extent to which a particular woman deviates from the median of a healthy population; it does not indicate the probability of suffering an adverse outcome at some time in the future. The latter requires empirical evidence on the sensitivity distribution, i.e. the probability of suffering a given adverse outcome, which is in turn dependent on the prevalence of the outcome. As shown in Fig. 17, this distribution may deviate from the specificity distribution by a large or small amount, reflecting the steepness of the rise in risk as deviation from the median increases. The quantitative relationship between the two distributions (i.e. the distance between them) can only be determined empirically.

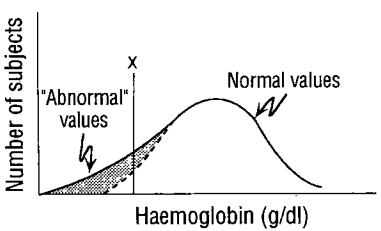
The importance of this lies in the fact that the "conventional" cut-off points based on the specificity distribution (e.g. 3rd percentile) have no intrinsic value for predicting the risk of adverse outcomes such as SGA, preterm delivery, obstetric complications, neonatal or maternal mortality. In general, it is safe to assume that the risk of a given adverse outcome increases at some point as deviation from the median of the specificity distribution increases, but the location of that point and the steepness of the increase in risk beyond that point cannot be predicted without empirical support. It follows that the most efficient screening cut-off point for predicting an outcome cannot be identified by using only the specificity distribution; knowledge of the sensitivity distribution and the prevalence of the outcome is also necessary.

As shown in Fig. 17, similar concepts apply to the development of indicators of benefit. In this case it would be desirable to compare the distributions of those who benefit from a given intervention and those who derive little or no benefit with respect to a measurement taken before intervention. For example, the mid-pregnancy weight of women who benefited from supplementation could be compared with the weight of those who did not. ("Benefit" might be defined as a higher birth weight or postpartum maternal BMI than was predicted on the basis of other characteristics, e.g. height, MUAC, previous LBW, poor socioeconomic status.) With this information it would be possible to identify the most efficient cut-off point for identifying women likely to benefit from supplementation. Note that the definition of "benefit" can be based on the same measurements as those used to predict who will benefit or on different measurements; in this example, benefit might be defined relative to SGA, or to maternal weight at some point after the intervention.

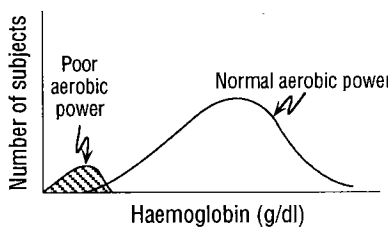
Finally, indicators of response refer to characteristics that are capable of changing in response to a given intervention. In the above example, indicators of response may be fetal growth, maternal anthropometry during or after intervention, or, less obviously, such characteristics as

Figure 17
Characteristics of an indicator that predicts benefit from an intervention[a]

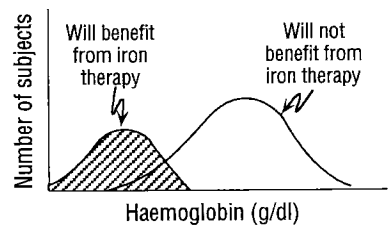
A. Normative approach
X is the cut-off point below which an individual is considered to have an undesirable condition (e.g. anaemia)



B. Risk approach
The shaded area indicates those whose performance (measured here as oxygen intake or "aerobic power") is below their potential



C. Benefit approach
The shaded area indicates those who could benefit from an intervention. The proportion of the total population represented by this group may be different from that in A or B



[a] Reproduced, with permission, from reference *51*.

maternal physical activity. It is important to note that indicators of response need not be the same as indicators of risk or predictors of benefit; indeed, some indicators are intrinsically more responsive to interventions in ways that do not necessarily conform to *a priori* expectations. Beaton & Ghassemi (*34*) suggested that physical activity might be a more appropriate indicator of response to supplementary feeding in children than anthropometric indicators, even though the latter may be useful in predicting who will benefit and are also good indicators of risk of adverse outcomes such as mortality. Similarly, Wolgemuth et al. (*52*) noted that dietary supplementation of road workers in Kenya did not result in greater work productivity, but anecdotal evidence suggested that workers were more active in the domestic sphere. It has also been noted that height of infants or young children is more responsive to supplementary feeding than is arm circumference or skinfold thickness (*53*). As illustrated by these examples, the responsiveness of indicators is a particularly important attribute for evaluation of programme impact.

In contrast to indicators of risk and benefit, there is no particular need for reference data for indicators of response. Instead, it has been suggested that responsiveness should be viewed as a continuous variable, such that different indicators may be characterized as being more or less responsive than others (*53*). The proposed formula for responsiveness is $0.5\ (\text{response/SD})^2$, where response is the difference between the mean value of the indicator in treated and control groups and SD is the pooled standard deviation of the two groups before treatment. As more is learned about the responsiveness of various indicators to different interventions, it will be possible to select more carefully those that are appropriate in a given situation.

It should be stressed that empirical evidence is not currently sufficient for constructing reference data based on risk or benefit, nor is it adequate for guiding the selection of the most efficient indicators of response. These are priority areas for future research, especially as regards the development of predictors of benefit. This can be addressed through careful design of evaluation components in small- or large-scale intervention programmes and appropriate analysis of data. In time, such research would improve the targeting, screening, and evaluation of those same programmes and would assist in similar activities in programmes implemented elsewhere. Thus, it is suggested that these gaps in knowledge be addressed in the context of action programmes rather than separate research projects. Although this will require some increased investment in evaluation and analysis, the cost-effectiveness of such an approach is likely to be far greater than that of separate research.

3.8 Populations for which compiled reference data are not pertinent

Since the only compiled reference data currently in use are from developed countries, there is some question of their usefulness in less developed areas. Total gestational weight gains are generally 4–5 kg

greater in advantaged than in disadvantaged or undernourished populations. Moreover, when the two types of population are compared on the basis of similar birth weights and gestational ages, women from developed countries apparently gain about 3–4 kg more than those in less developed areas who produce infants of the same size (Fig. 14), suggesting that reference data for weight gain (either total or incremental) derived from developed countries may be excessive. This requires further investigation in relation to other pregnancy outcomes, including maternal body stores necessary to support lactation and prevent postpartum nutritional depletion and its sequelae.

The WHO Collaborative Study (*23*) provided reasonable evidence to suggest that, for pregnancy risk, there are different absolute cut-off values of maternal height for different populations. This should be investigated more thoroughly, using the WHO collaborative data with different cut-off values for each population represented in the data set.

For other measures, such as arm and calf circumference, there is insufficient evidence in the literature to allow different reference data to be recommended for different populations. Only after better information is available on the relationship of these indicators to specific outcomes can the question of reference data be considered.

Current reference data do not take into consideration potential modifying factors that may influence the interpretation of risk in relation to cut-offs. These factors include extremes in maternal age and parity, and pre-existing or current diseases such as diabetes, malaria, and anaemia.

3.9 The use and interpretation of anthropometry in lactating women

3.9.1 *Biological significance of anthropometry during lactation*

Changes in weight and body composition that occur during lactation underlie the biological basis for anthropometric assessment of lactating women, and the same rationale that governs the use of anthropometric measurements as indicators of nutritional risk or predictors of benefit from intervention applies to lactating women. A lack of reference data, however, limits the application of anthropometry to lactating women.

Lactation is the most energetically demanding phase of the human reproductive cycle. The total energy cost of producing milk is estimated to be 2930 kJ/day during the first 6 months of lactation and 2090 kJ/day during the next 18 months (*54*). Normally, fat deposited during pregnancy (about 4 kg, but this is highly variable) is mobilized postpartum to meet the energy costs of lactation (*55*).

Fat seems to be deposited preferentially in pregnancy, notably in the back and upper thighs, but not over the arms. This same pattern has been observed in different populations (*6, 56*). However, the increase of skinfold thicknesses cannot be attributed solely to deposition of fat,

since, in most body sites, it is followed by a decrease at the time of parturition (*6*). During lactation lipolysis is higher in the femoral than in the abdominal region.

Thereafter, weight loss is slow and stabilizes at about 4–6 months. This is variable, however, and depends on socioeconomic status, weight gained during pregnancy, energy intake, and pattern of breast-feeding. It is useful to summarize the weight changes that occur at two levels of nutritional status. In well nourished lactating women, changes are generally minor and gradual. Weight losses are highest in the first 3 months of lactation (*57–61*) and are generally reported to be greater in women who breast-feed exclusively (*58–60, 62*). Skinfold thickness also tends to reflect weight changes, with most measurement sites showing decreased thickness as lactation progresses (*56–59, 63, 64*). An exception is the apparent gain in triceps skinfold thickness reported by several authors (*56, 58, 59, 63–65*). Among undernourished women, gestational weight gain and postpartum weight loss are lower than in well nourished women (*6, 66–75*). Although published values for milk composition of women from developing countries differ substantially, lower nutrient levels have usually been found in undernourished women (*57, 62, 68, 76*).

3.9.2 *Selection of individuals*

At present, anthropometric measurements cannot be used effectively to assess the nutritional status of individual lactating women. Moreover, no anthropometric indicators of risk for undesirable outcomes or of benefit from medical or nutritional interventions have been developed specifically for lactating women (*77*).

Nutritional status of the mother during lactation depends on many factors such as past nutritional status, weight gain in pregnancy, immediate postpartum weight loss, duration and intensity of lactation, dietary intake, and physical activity. Studies conducted worldwide have consistently noted that weight loss during lactation is much greater in the first month because of the shedding of extra water, tissue, and, to some extent, fat accrued during pregnancy.

While the limited literature reviewed here on changes in maternal weight is useful for establishing a basis for anthropometric assessment of nutritional status during lactation, it is only a first step towards developing anthropometric indicators. The purpose for which these indicators will be used must be considered. If it is for screening women at risk of poor postpartum outcomes, more extensive information on the nature of these outcomes and how they are affected by maternal nutrition will be required. Knowledge in this area is very limited at present and the research conducted thus far has not reported results in a manner that is easily interpreted for evaluation of indicators of risk. Even the definition of desirable outcomes presents problems. Lactation performance is certainly one area for investigation; however, the difficulties associated

with evaluating the quantity and quality of breast milk as well as the definition of optimal growth in breast-fed infants (see section 5) suggest that more research is needed before progress can be made with identification and validation of indicators of risk for lactating women. Analysis of outcomes should also consider mothers' health and well-being during and after lactation. These considerations include resumption of menstruation, depletion and repletion of nutritional stores, and development of a risk profile for various chronic diseases.

3.9.3 *Characteristics for the development of normative reference data*

Normative reference data are not available to identify nutritionally "at risk" groups of lactating women (*76*). In populations in which anthropometric indices reflect food availability, nutritional vulnerability should be indicated by body weight, skinfold thicknesses, and arm and calf circumferences. Poor gestational weight gain may predict poor lactation performance, because fat stores may be inadequate to subsidize the energy costs of lactation. Although extensive data are available on the milk production of women of varying nutritional status, they have not been used to develop indicators of lactation performance.

Since no reference data exist for assessing nutritional status during lactation and very little research has been conducted in this area, only provisional criteria can be recommended for lactating women. There is evidence that poor maternal postpartum status, reflected in low BMI, is associated with poor lactation performance and poor infant growth, which suggests that BMI may be a useful indicator of postpartum nutritional status. However, the level of BMI below which there is a risk of poor lactation or infant growth has not been reported. It is possible to estimate a level based on the lower limit of BMI (<18.5) suggested for thin adults in section 8, adjusted for the average weight (4 kg) retained by mothers following an acceptable pregnancy weight gain (10.5-12.0 kg) and enough time for postpartum hydration to have equilibrated (2-4 weeks). This results in an estimated cut-off for BMI of 20.3 at 1 month postpartum for women 150 cm tall. BMI may be expected to decline steadily throughout the first 6 months of lactation, at which point the non-pregnant non-lactating value of 18.5 can be used as a cut-off for identifying women at risk.

Only a limited number of studies have attempted to assess upper levels of BMI during lactation. However, in the light of the recommendations for modest gestational weight gain by overweight and obese women (*9*), it is likely that the upper limits of BMI recommended for non-pregnant, non-lactating women (see section 7) would apply to lactating women as well.

To develop normative reference data for individuals, anthropometric measurements would have to be recorded longitudinally in a population of well nourished, healthy, lactating women and related to their lactation performance. Anthropometric changes in lactating women have been

documented in a number of studies, but few of these have assessed lactation performance. Acceptable limits of postpartum weight and body compositional changes would be defined on the basis of lactation performance. In prolonged lactation, success would be at the expense of maternal stores. Development of anthropometric reference data for lactating women would require ancillary data on age, parity, prepregnant weight, gestational weight gain, and the intensity (exclusive or partial) and duration of breast-feeding. An evaluation of lactation performance would require data on milk volume, milk composition, and growth in infant weight and length.

Evidence to date does not suggest an association between maternal anthropometric indices and early lactation performance in well nourished populations. In prolonged lactation, maternal adipose stores may limit lactation performance if dietary intake is restricted.

Normative reference data based on anthropometric changes in well nourished populations are unlikely to be applicable to lactating women in undernourished populations because of significant differences in height, weight, and gestational weight gain between the populations. Moreover, sensitive indicators may be population-specific; for example, the triceps skinfold thickness is indicative of milk fat concentration in Bangladeshi and Gambian women, but not in American women (*67, 68*).

Evidence of an association between poor nutritional status and compromised lactation performance supports the development of anthropometric indicators within populations of nutritionally vulnerable lactating women. In undernourished populations, anthropometry will reflect both past and present food availability. Critical anthropometric thresholds should be definable, below which restricted maternal diet and limited tissue reserves are inadequate to meet the energy demands of lactation.

3.9.4 *Research needs for lactating women*

Reference data are needed for estimation of the prevalence of undernutrition among lactating women in the population, and can be developed on the basis of available data from well nourished women. This should take account of different patterns of weight gain during pregnancy and different breast-feeding patterns. The same reference data would be used to screen individual lactating women for interventions.

To predict risk of maternal malnutrition or of individual women producing insufficient milk to maintain normal infant growth, there is a need for risk indicators. No indicators of risk of adverse maternal or infant outcomes are available at present, and research on their development is essential. Candidate indicators include maternal body weight or its change over a short period of time, maternal calf circumference, change in maternal skinfold thickness over a short period of time, and poor infant growth during exclusive breast-feeding.

Indicators are also needed that will predict benefits to the individual lactating woman or her breast-feeding infant of an appropriate intervention. Candidate indicators include maternal body weight or change in weight over a short period, maternal skinfold thicknesses or change in thicknesses over a short period of time, infant milk intake, and the ability to maintain exclusive breast-feeding of infants up to 6 months of age.

Candidate indicators for evaluating the response of individual lactating women to an appropriate intervention include changes in maternal body weight and skinfold thicknesses, change in infant milk intake, and the proportion of women who are able to breast-feed exclusively. Most of these indicators have been used in a recent randomized study of nutritional intervention in Guatemala (*78*).

3.10 Conclusions and recommendations

3.10.1 *For practical implementation*

Conclusions

Anthropometry in some form will continue to be a routine part of prenatal examinations throughout the world. This report has identified several applications of anthropometry that are useful in specific circumstances, depending on the availability of resources and the potential for intervention to achieve favourable pregnancy outcomes. The criterion of utility for most anthropometry examined in the report is a degree of association between the anthropometric indicator and the risk of a specific undesirable outcome such as SGA, preterm delivery, delivery complications, and, to a lesser degree, postpartum maternal depletion. In very few instances have the sensitivity, specificity, and positive predictive value of these relationships been examined to test for misclassification of individuals for risk of poor outcome or response to interventions. Any recommendations for the use of specific anthropometry summarized in Tables 9 and 10 are therefore provisional. The analysis of misclassification undertaken by WHO (*23*) is, conceptually, an appropriate next step in the evaluation of indicators that show significant association with pregnancy outcomes. However, there is a need for further analysis of sensitivity, specificity, and positive predictive value at different cut-offs for the various indicators. Within these limitations, it is possible to rank the various indicators, as applied to individuals in clinical settings, according to their ORs (Table 3), sensitivities, and specificities (Table 4) for different outcomes and with different levels of logistic support.

When resources are limited (i.e. no scales are available), short height may be useful as a screening instrument owing to modest ORs (1.2 to 1.9) for several outcomes. However, it is likely that the use of height will result

in considerable misclassification. The appropriate height cut-off for screening will depend on local conditions, such as resources for intervention. Calf circumference is a more promising indicator but must be investigated in several different settings. Thus far, only maternal height seems to be a reliable predictor of the need for assisted delivery; research on its relationship to maternal mortality is needed, especially among disadvantaged populations.

When scales are available and properly used, attained body weight at any time during pregnancy appears to be the most useful screening indicator for SGA. However, further testing of the potential for misclassification is needed and, if the results are sufficiently promising, proper reference data must be developed. Weight measurements should be taken early enough in pregnancy to allow appropriate intervention. Attained weight at 20 weeks of gestation is useful for screening for dietary supplementation, although earlier assessment is preferred. Later assessment can be useful for referring mothers for delivery to a health facility where SGA and preterm infants can receive special care. The utility of weight measured during pregnancy is improved if short height is used as a first level of screening; body mass index alone is less useful than this two-stage screening approach, and further analysis of misclassification in two-stage screening is needed. Calf circumference may also prove more useful than some of the more "sophisticated" indicators. Weight gained between two examinations is less useful than a single measurement of weight as a predictor of poor pregnancy outcome. Change in thigh skinfold thickness appears to have greater potential as a specific indicator of maternal body composition change, and is related to fetal growth and postpartum maternal body stores.

Recommendations to countries

1. To screen women for risk of delivering SGA babies when scales are unavailable, height should be measured at any stage of pregnancy and deviations from local norms should be interpreted with regard to resources available for supporting interventions.
2. When scales are available, women should be weighed as early in pregnancy as possible to screen for risk of SGA babies. Maternal weight for gestational age should be compared with a reference similar to that shown in Fig. 15 for advantaged populations. Until more appropriate references are available, the same reference may be used for less well nourished populations, but cut-off levels for screening individuals should be set lower and should take account of local resource availability.
3. For monitoring the response of individuals in all populations to interventions during pregnancy, weight changes between successive examination should be determined during the final two trimesters and compared with existing cut-off guidelines set by the Institute of Medicine (*9*).

4. Since the application and interpretation of anthropometry during pregnancy are likely to differ from country to country, each country should develop its own cut-off values for each relevant indicator, using methodologies described in this report.
5. Countries should develop nutritional surveillance systems that cover the collection and analysis of data on national problems of pregnancy and lactation, and appropriate action.

3.10.2 *For future research and the collection of reference data*

1. Many of the practical recommendations in this report are based on the composite analysis of 25 studies reported by the WHO Collaborative Study (*23*). Further analysis of these data is required to establish the precise ORs for different anthropometry versus different pregnancy outcomes. Research of this type should also be undertaken using other data sets not included in the WHO study in order to broaden knowledge in this area; this should include analysis of sensitivity, specificity, and positive predictive value of various indicators and cut-off points for different pregnancy outcomes in relation to specific purposes and the resources available in different settings.
2. Most of the pregnancy outcomes used for the evaluation of maternal anthropometry in this report are commonly reported in the research literature, but other important outcomes in the areas of fetal, neonatal, and maternal morbidity and mortality, as well as fertility, should also be tested. More research is urgently needed on maternal postpartum consequences of anthropometric variation during pregnancy, especially in relation to lactation performance, resumption of ovarian function, and repletion (or further depletion) of maternal nutrient stores. Some of the common outcome measures should be refined. Classifications such as SGA, preterm and assisted delivery, and maternal morbidity are themselves subject to further subclassification, and the resulting subgroups are likely to be more homogeneous with regard to both causes and consequences; maternal anthropometry may be more effective as an indicator of risk for some of the more specific outcomes and unrelated to risk of others.
3. Studies of anthropometric risk factors for poor pregnancy outcomes will also benefit by expansion of the analysis to include both long-term outcomes (infant death, poor postnatal cognitive development) and important intermediate measures (SGA, risk of preterm delivery, breast milk quantity and quality). This would allow more precise causal pathways to be identified and the validity of using commonly measured intermediate variables as outcomes to be established. In general, research is needed to establish the biological bases of the anthropometric correlates of specific pregnancy outcomes. For example: What is short maternal stature actually measuring as it relates to increased risk of SGA? What proportion of maternal weight

gain is represented by maternal tissues that will support the pregnancy or later lactation?

4. Many of the newer applications of anthropometry have been validated in only a few isolated situations. Validation of calf circumference and thigh skinfold measurements should be extended to different populations and settings. The recommended weight gain chart of Rosso (*30*) should be tested for screening efficiency in populations where maternal malnutrition is prevalent, and the algorithm proposed by WHO to screen women for supplementation by using a single weight measurement from the second half of pregnancy should be evaluated. Systematic evaluation of multi-stage screening approaches that employ a cascade, or sequence, of measurements is also needed.
5. The comprehensive treatment of maternal anthropometry of undernourished population groups should be extended to populations where excess energy balance is more common. It is also important to consider the effects of maternal obesity and high weight gain on such pregnancy outcomes as SGA, macrosomia, preterm delivery, complications of labour and delivery, postpartum maternal weight retention, and infant and maternal morbidity. Some research recommendations in this area were made by the Institute of Medicine report (*9*), which focused primarily on gestational weight gain of women from more developed countries, and more results are starting to appear in the literature (*79*). An extensive list of research recommendations for women from less developed countries has already been compiled by Krasovec & Anderson (*10*).

Completion of even part of this research agenda will help to establish criteria for reference data for assessing the nutritional status of women during pregnancy. Continued analysis of data sets from the WHO Collaborative Study (*23*) for misclassification relative to various pregnancy outcomes is particularly important. Choice of indicators and cut-off values requires an analysis of indicators relative to commonly defined outcomes across a variety of clinical and public health settings in several different population groups that reflect geographical, demographic, ethnic, and socioeconomic variation.

3.10.3 *For WHO*

1. WHO should facilitate research on anthropometry of women during the reproductive years, based especially on the data from the WHO Collaborative Study (*23*). Analysis of these data should be extended to include different population groups and various maternal and infant postpartum outcomes.
2. WHO should assist in the development of methods that will allow countries to establish locally relevant cut-off values for the anthropometric indicators recommended in this report. Data from the WHO Collaborative Study could serve as a useful resource in this effort.

4. WHO should facilitate the development of simple algorithms for the application and interpretation of anthropometry during pregnancy and lactation, and for integrating anthropometry into health care strategies.
5. WHO should assist countries in establishing surveillance systems to identify solutions to health and nutrition problems in women of reproductive age.

References

1. Villar J et al. Perinatal data reliability in a large teaching obstetric unit. *British journal of obstetrics and gynaecology*, 1988, **95**:841-848.
2. Dawes MG, Green J, Ashurst H. Routine weighing in pregnancy. *British medical journal*, 1992, **304**:487-489.
3. Merchant K, Martorell R. Frequent reproductive cycling: does it lead to nutritional depletion of mothers? *Progress in food and nutrition science*, 1988, **12**:339-369.
4. Kramer MS et al. Maternal nutrition and spontaneous preterm birth. *American journal of epidemiology*, 1992, **136**:574-583.
5. Villar J, Belizan JM. The evaluation of the methods used in the diagnosis of intrauterine growth retardation. *Obstetrical and gynecological surveys*, 1986, **41**:187-199.
6. Villar J et al. Effect of fat and fat-free mass deposition during pregnancy on birth weight. *American journal of obstetrics and gynecology*, 1992, **167**:1344-1352.
7. Savitz DA, Blackmore CA, Thorp JM. Epidemiologic characteristics of preterm delivery: etiologic heterogeneity. *American journal of obstetrics and gynecology*, 1991, **164**:467-471.
8. Kramer MS et al. The validity of gestational age estimation by menstrual dating in term, preterm, and postterm gestations. *Journal of the American Medical Association*, 1988, **260**:3306-3308.
9. Institute of Medicine. *Nutrition during pregnancy. Part 1: Weight gain*. Washington, DC, National Academy Press, 1980.
10. Krasovec K, Anderson MA, eds. *Maternal nutrition and pregnancy outcomes: anthropometric assessment*. Washington, DC, Pan American Health Organization, 1991 (Scientific Publication, No. 529).
11. Stevens-Simon C, McAnarney ER, Coulter MP. How accurately do adolescents estimate their weight prior to pregnancy? *Journal of adolescent health care*, 1986, **7**:250-254.
12. Wilcox AJ, Horney LE. Accuracy of spontaneous abortion recall. *American journal of epidemiology*, 1984, **120**:727-733.
13. Launer L, Villar J, Kestler E. Epidemiological differences among birthweight and gestational age subgroups of newborns. *American journal of human biology*, 1991, **3**:425-433.
14. Hytten FE. Weight gain in pregnancy. In: Hytten FE, Chamberlain G, eds. *Clinical physiology in obstetrics*. Oxford, Blackwell, 1980:193-233.

15. Lawrence M, McKillop FM, Durnin JVG. Women who gain more fat during pregnancy may not have bigger babies: implications for recommended weight gain during pregnancy. *British journal of obstetrics and gynaecology*, 1991, 98:254-259.

16. Martorell R, Habicht J-P. Growth in early childhood in developing countries. In: Falkner F, Tanner JM, eds. *Human growth: a comprehensive treatise*, 2nd ed., Vol. 3. New York, Plenum Press, 1986:241-262.

17. Scholl TO, Hediger ML, Ances IG. Maternal growth during pregnancy and decreased infant birth weight. *American journal of clinical nutrition*, 1990, 51:790-793.

18. González-Cossío T, Flores F, ARCIU Group. Validity of maternal calf circumference to identify risk of intrauterine growth retardation (IUGR). *The FASEB journal*, 1992, 6(5):A1683.

19. McDonald E. Mensuration of the child in the uterus with new methods. *Journal of the American Medical Association*, 1906, 47:1979-1983.

20. Jacobsen G. Prediction of fetal growth deviations by use of symphysis-fundus height measurements. *International journal of technology assessment in health care*, 1992, 8(Suppl. 1):152-159.

21. Hughes AB et al. Symphysis-fundus height, maternal height, labor pattern, and mode of delivery. *American journal of obstetrics and gynecology*, 1987, 156: 644-648.

22. Beaton GH et al. *Appropriate uses of anthropometric indices in children : a report based on an ACC/SCN workshop.* New York, United Nations Administrative Committee on Coordination/Subcommittee on Nutrition, 1990 (ACC/SCN State-of-the-Art Series, Nutrition Policy Discussion Paper No. 7).

23. Maternal anthropometry and pregnancy outcomes: a WHO collaborative project. *Bulletin of the World Health Organization*, 1995, 73 (Suppl.).

24. Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of the World Health Organization*, 1987, 65:663-737.

25. Habicht J-P. Some characteristics of indicators of nutritional status for use in screening and surveillance. *American journal of clinical nutrition*, 1980, 33: 531-535.

26. Galen RS, Gambino SR. *Beyond normality: the predictive value and efficacy of medical diagnoses.* New York, Wiley, 1975.

27. Kalkwarf HJ. *Maternal weight gain during pregnancy and risk of preterm delivery: effects on neonatal mortality and public health impact* (Dissertation). Ithaca, NY, Cornell University, 1991.

28. Lechtig A. Predicting risk of delivering low birth weight babies: which indicator is better? *Journal of tropical pediatrics*, 1988, 34:34-41.

29. Merchant KM, Villar J. Effect of maternal supplementation on risk of perinatal distress and intrapartum cesarean delivery. *The FASEB journal*, 1993, 7:A282.

30. Rosso P. A new chart to monitor weight during pregnancy. *American journal of clinical nutrition*, 1985, 41:644-652.

31. **Haas JD.** Weight gain in pregnancy. In: Krasovec K, Anderson MA, eds. *Maternal nutrition and pregnancy outcomes: anthropometric assessment*. Washington, DC, Pan American Health Organization, 1991 (Scientific Publication, No. 529): 29-51.

32. **Belizan JM et al.** Diagnosis of intrauterine growth retardation by a simple clinical method: measurement of uterine height. *American journal of obstetrics and gynecology*, 1978, **131**:643-646.

33. **Susser M.** Maternal weight gain, infant birth weight, and diet: causal sequences. *American journal of clinical nutrition*, 1991, **53**:1384-1396.

34. **Beaton GH, Ghassemi H.** Supplementary feeding programs for young children in developing countries. *American journal of clinical nutrition*, 1982, **35**(4, suppl.): 863-916.

35. **Habicht J-P, Butz WP.** Measurements of health and nutrition effects of large-scale nutrition intervention projects. In: Klein RE et al., eds. *Evaluating the impact of nutrition and health programs*. New York, Plenum Press, 1979:133-182.

36. **Valverde V et al.** Uses and constraints of school children's height data for planning purposes: national experiences in Central America. *Food and nutrition bulletin*, 1986, **8**(3):42-48.

37. **Lubchenco LO, Searls DT, Brazie JV.** Neonatal mortality rate: relationship to birth weight and gestational age. *Journal of pediatrics*, 1972, **81**:814-822.

38. **Haas JD, Balcazar H, Caulfield I.** Variation in early neonatal mortality for different types of fetal growth retardation. *American journal of physical anthropology*, 1987, **73**:467-473.

39. **Garner P, Kramer MS, Chalmers I.** Might efforts to increase birthweight in undernourished women do more harm than good? *Lancet*, 1992, **340**:1021-1023.

40. **Mason JP et al.** *Nutritional surveillance*. Geneva, World Health Organization, 1984.

41. **Abraham S, Johnson CL, Najjar MF.** *Weight by height and age for adults 18-74 years*. Hyattsville, MD, US Department of Health, Education, and Welfare, 1979 (Vital and Health Statistics: Data from the National Health Survey, Series 11, No. 208).

42. **Eveleth PB, Tanner JM.** *Worldwide variation in human growth*. Cambridge, Cambridge University Press, 1990.

43. **Metropolitan Life Insurance Company.** New weight standards for men and women. *Statistics bulletin*, 1959, **40**:1-4.

44. **Tompkins WT, Wiehl DG.** Nutritional deficiencies as a causal factor in toxemia and premature labor. *American journal of obstetrics and gynecology*, 1951, **62**:898-919.

45. **Thomson AM, Billewicz WZ.** Clinical significance of weight trends during pregnancy. *British medical journal*, 1957, **1**:243-247.

46. **Tanner JM.** *Growth at adolescence; with a general consideration of the effects of hereditary and environmental factors upon growth and maturation from birth to maturity*, 2nd ed. Oxford, Blackwell, 1962.

47. **Hediger ML et al.** Patterns of weight gain in adolescent pregnancy: effects on birth weight and preterm delivery. *Obstetrics and gynecology*, 1989, **74**:6-12.

48. Cameron N. Conditional standards for growth in height of British children from 5.0 to 15.99 years of age. *Annals of human biology*, 1980, 7:331-337.

49. Naeye RL. Weight gain and the outcome of pregnancy. *American journal of obstetrics and gynecology*, 1979, **135**:3-9.

50. Habicht J-P, Meyers LD, Brownie C. Indicators for identifying and counting the improperly nourished. *American journal of clinical nutrition*, 1982, **35**(5, suppl.): 1241-1254.

51. Rasmussen K, Habicht J-P. Malnutrition among women: indicators to estimate prevalence. *Food and nutrition bulletin*, 1989, 11:29-37.

52. Wolgemuth JC et al. Worker productivity and the nutritional status of Kenyan road construction laborers. *American journal of clinical nutrition*, 1982, **36**:68-78.

53. Habicht J-P, Mason J, Tabatabi H. Basic concepts for the design of evaluation during programme implementation. In: Sahn DE, Lockwood R, Scrimshaw NS, eds. *Methods for the evaluation of the impact of food and nutrition programs*. Tokyo, United Nations University, 1984 (UNU Food and Nutrition Bulletin, Supplement No. 8).

54. *Energy and protein requirements. Report of a Joint FAO/WHO/UNU Expert Consultation*. Geneva, World Health Organization, 1985 (WHO Technical Report Series, No. 724).

55. Hytten FE. Nutrition. In: Hytten FE, Chamberlain G, eds. *Clinical physiology in obstetrics*. Oxford, Blackwell, 1980.

56. Taggart NR et al. Changes in skinfolds during pregnancy. *British journal of nutrition*, 1967, 21:439-451.

57. Butte NF et al. Effect of maternal diet and body composition on lactational performance. *American journal of clinical nutrition*, 1984, **39**:296-306.

58. Manning-Dalton C, Allen LH. The effects of lactation on energy and protein consumption, postpartum weight change and body composition of well-nourished North American women. *Nutrition research*, 1983, 3:293-308.

59. Brewer MM, Bates MR, Vannoy LP. Postpartum changes in maternal weight and body fat deposits in lactating vs nonlactating women. *American journal of clinical nutrition*, 1989, **49**:259-265.

60. Ohlin A, Rossner S. Maternal body weight development after pregnancy. *International journal of obesity*, 1990, 14:159-173.

61. Morse EH et al. Comparison of the nutritional status of pregnant adolescents with adult pregnant women. II. Anthropometric and dietary findings. *American journal of clinical nutrition*, 1975, 28:1422-1428.

62. Paul AA, Müller EM, Whitehead RG. Seasonal variations in energy intake, body-weight and skinfold thickness in pregnant and lactating women in rural Gambia. *Proceedings of the Nutrition Society*, 1979, **38**:28A.

63. Forsum E, Sadurskis A, Wager J. Estimation of body fat in healthy Swedish women during pregnancy and lactation. *American journal of clinical nutrition*, 1989, **50**:465-473.

64. Forsum E, Sadurskis A, Wager J. Resting metabolic rate and body composition of healthy Swedish women during pregnancy. *American journal of clinical nutrition*, 1988, 47:942-947.

65. **Dugdale AE, Eaton-Evans J.** The effect of lactation and other factors on post-partum changes in body-weight and triceps skinfold thickness. *British journal of nutrition*, 1989, **61**:149-153.

66. **Kramer FM et al.** Breastfeeding reduces maternal lower body fat. *Journal of the American Dietetic Association*, 1993, **93**:429-443.

67. **Prentice A, Prentice AM, Whitehead RG.** Breast-milk fat concentrations of rural African women. 2. Long-term variations within a community. *British journal of nutrition*, 1981, **45**:495-503.

68. **Brown KH et al.** Lactational capacity of marginally nourished mothers: relationships between maternal nutritional status and quantity and proximate composition of milk. *Pediatrics*, 1986, **78**:909-919.

69. **Schutz Y, Lechtig A, Bradfield RB.** Energy expenditures and food intakes of lactating women in Guatemala. *American journal of clinical nutrition*, 1980, **33**:892-902.

70. **Ebrahim GJ.** *Cross-cultural aspects of breast-feeding. Breast-feeding and the mother*. Amsterdam, Elsevier, 1976 (Ciba Foundation Symposium, No. 45): 195-204.

71. **Adair LS, Pollitt E.** Seasonal variation in pre- and postpartum maternal body measurements and infant birth weights. *American journal of physical anthropology*, 1983, **62**:325-331.

72. **Adair LS, Pollitt E.** Outcome of maternal nutritional supplementation: a comprehensive review of the Bacon Chow study. *American journal of clinical nutrition*, 1985, **41**:948-978.

73. **Adair LS, Pollitt E, Mueller WH.** Maternal anthropometric changes during pregnancy and lactation in a rural Taiwanese population. *Human biology*, 1983, **55**:771-787.

74. **Adair LS, Pollitt E, Mueller WH,** The Bacon Chow study: effect of nutritional supplementation on maternal weight and skinfold thickness during pregnancy and lactation. *British journal of nutrition*, 1984, **51**:357-369.

75. **Calloway DH, Murphy SP, Beaton GH.** *Food intake and human function: a cross-project perspective*. Berkeley, CA, University of California, 1988.

76. **Rasmussen KM.** Maternal nutritional status and lactational performance. *Clinical nutrition*, 1988, **7**:147-155.

77. **Institute of Medicine (U.S.) Subcommittee on Nutrition during Lactation.** *Nutrition during lactation*. Washington, DC, National Academy Press, 1991.

78. **González-Cossío T et al.** Food supplementation during lactation increases infant milk intake and proportion of exclusive breastfeeding. *The FASEB journal*, 1991, **5**:A917.

79. **Johnson JWC, Longmate JA, Frentzen B.** Excessive maternal weight and pregnancy outcome. *American journal of obstetrics and gynecology*, 1992, **167**:353-372.