

*This report contains the collective views of an international group of experts and does not necessarily represent the decisions or the stated policy of the World Health Organization*

**WHO Technical Report Series**

**854**

---

# **PHYSICAL STATUS: THE USE AND INTERPRETATION OF ANTHROPOMETRY**

---

Report of a  
WHO Expert Committee



---

**World Health Organization**

**Geneva 1995**

WHO Library Cataloguing in Publication Data

WHO Expert Committee on Physical Status : the Use and Interpretation of Anthropometry  
Physical status : the use and interpretation of anthropometry : report of a WHO expert  
committee.

(WHO technical report series ; 854)

1. Anthropometry 2. Health status indicators 3. Nutrition assessment I. Title II. Series

ISBN 92 4 120854 6  
ISSN 0512-3054

(NLM Classification: GN 54)

The World Health Organization welcomes requests for permission to reproduce or translate its publications, in part or in full. Applications and enquiries should be addressed to the Office of Publications, World Health Organization, Geneva, Switzerland, which will be glad to provide the latest information on any changes made to the text, plans for new editions, and reprints and translations already available.

**© World Health Organization 1995**

Publications of the World Health Organization enjoy copyright protection in accordance with the provisions of Protocol 2 of the Universal Copyright Convention. All rights reserved.

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

**Printed in Switzerland**

95/10429 – Benteli – 7000

# Contents

1. Introduction	1
References	3
2. Technical framework	4
2.1 Introduction	4
2.2 Levels of body composition	4
2.3 Anthropometric measurements, indices, and indicators	6
2.3.1 Measurements	6
2.3.2 Indices	7
2.3.3 Indicators	8
2.4 Selection of anthropometric indicators	10
2.5 Sensitivity and specificity of indicators	12
2.6 Selection of a best indicator	14
2.7 Using anthropometry in individuals	16
2.7.1 Screening with one measurement for targeting an intervention	16
2.7.2 Assessing response to an intervention	21
2.8 Using anthropometry in populations	22
2.8.1 Uses related to decisions	22
2.8.2 Targeting interventions	24
2.8.3 Assessing response to an intervention	25
2.8.4 Ascertaining the determinants and consequences of malnutrition	26
2.8.5 Nutritional surveillance	26
2.9 Characteristics of reference data	29
References	33
3. Pregnant and lactating women	37
3.1 Introduction	37
3.1.1 Background	37
3.1.2 Methodology	38
3.1.3 Biological significance of anthropometry during pregnancy	40
3.1.4 Anthropometry as an indicator of nutritional and health status	41
3.2 Using anthropometry in individuals	47
3.2.1 Choosing an indicator	47
3.2.2 Applications of anthropometry for screening pregnant women	56
3.2.3 Assessing response to an intervention	61
3.3 Using anthropometry in populations	70
3.3.1 Targeting interventions	72
3.3.2 Assessing response to an intervention	79
3.3.3 Ascertaining the determinants and consequences of malnutrition	80
3.3.4 Nutritional surveillance	81
3.4 Population data management and analysis	84
3.4.1 Sampling considerations	84
3.4.2 Problem identification	84
3.4.3 Policy and planning	85
3.4.4 Programme management and evaluation	86
3.5 Methods of taking measurements	87
3.6 Sources and characteristics of reference data	88
3.6.1 Existing reference data	88
3.6.2 Criteria for establishing reference data	98
3.6.3 Recommendations for new reference data	101
3.7 Relationship between normative reference data and functional outcomes	104

3.8	Populations for which compiled reference data are not pertinent	107
3.9	The use and interpretation of anthropometry in lactating women	108
3.9.1	Biological significance of anthropometry during lactation	108
3.9.2	Selection of individuals	109
3.9.3	Characteristics for the development of normative reference data	110
3.9.4	Research needs for lactating women	111
3.10	Conclusions and recommendations	112
3.10.1	For practical implementation	112
3.10.2	For future research and the collection of reference data	114
3.10.3	For WHO	115
	References	116
<b>4.</b>	<b>The newborn infant</b>	<b>121</b>
4.1	Introduction	121
4.2	Using anthropometry in individual newborn infants	123
4.3	Neonatal anthropometric assessment in populations	128
4.4	Selection of anthropometric indicators	129
4.4.1	Gestational age	129
4.4.2	Birth weight	135
4.4.3	Birth length	135
4.4.4	Birth head circumference	135
4.4.5	Proportionality indices	135
4.4.6	Other measurements	135
4.5	Reference data for size at birth	136
4.5.1	Criteria for evaluating existing references	136
4.5.2	Size at birth in early gestation	137
4.5.3	Size at birth in later gestation	138
4.6	Conclusions	149
4.7	Recommendations	153
4.7.1	General	153
4.7.2	For individuals	153
4.7.3	For populations	154
4.7.4	For WHO	154
4.7.5	For Member States	154
4.7.6	For future research	155
	References	155
<b>5.</b>	<b>Infants and children</b>	<b>161</b>
5.1	Introduction	161
5.1.1	Terminology and clarification of commonly used terms	162
5.1.2	Expression and interpretation of anthropometry	176
5.1.3	Biological and social significance of anthropometry	177
5.1.4	Issues in using anthropometric measurements as indicators of nutritional and health status	181
5.1.5	Conditioning the interpretation of anthropometry	182
5.2	Using anthropometry in individuals	183
5.2.1	Introduction	183
5.2.2	Screening children for health and nutritional disorders	183
5.3	Using anthropometry in populations	198
5.3.1	Introduction	198
5.3.2	Targeting interventions	198
5.3.3	Assessing response to an intervention	209
5.3.4	Ascertaining the determinants of malnutrition	210
5.3.5	Ascertaining the consequences of malnutrition	211
5.3.6	Nutritional surveillance	212

5.4	Population data management and analysis	215
5.4.1	Description of sources of data	215
5.4.2	Documentation and analysis of coverage rates	215
5.4.3	Reliability and validity	215
5.4.4	Data integrity or quality measures	217
5.4.5	Data compilation and documentation	219
5.4.6	Data analysis and presentation	219
5.5	Methods of measurement	224
5.5.1	Height measurements	224
5.5.2	Weight measurements	224
5.5.3	Age determination	224
5.6	Sources and characteristics of reference data	224
5.6.1	Issues related to selection and application of references	224
5.6.2	Local versus international references	225
5.6.3	Factors affecting the use and interpretation of growth references	226
5.6.4	The current international reference (NCHS/WHO reference)	227
5.6.5	Growth velocity curves	250
5.7	Presentation of anthropometric reference data	251
5.7.1	For individual-based applications	251
5.7.2	For population-based applications	252
5.8	Recommendations	253
5.8.1	Infants	253
5.8.2	Children	253
	References	255
<b>6.</b>	<b>Adolescents</b>	<b>263</b>
6.1	Introduction	263
6.1.1	Background	263
6.1.2	Biological and social significance of anthropometry	265
6.1.3	Anthropometry as an indicator of nutritional and health status	270
6.1.4	Conditioning the interpretation of anthropometry	276
6.2	Using anthropometry in individuals	281
6.2.1	Introduction	281
6.2.2	Screening for interventions	281
6.2.3	Assessing response to an intervention	283
6.3	Using anthropometry in populations	288
6.3.1	Introduction	288
6.3.2	Targeting interventions	288
6.3.3	Assessing response to an intervention	289
6.3.4	Ascertaining determinants of malnutrition	290
6.3.5	Ascertaining consequences of malnutrition	291
6.3.6	Nutritional surveillance	292
6.4	Population data management and analysis	300
6.4.1	Description of sources of data	300
6.4.2	Documentation and analysis of coverage rates	301
6.4.3	Documentation and analysis of reliability	301
6.4.4	Data compilation and documentation	301
6.4.5	Data analysis and presentation	302
6.5	Sources and characteristics of reference data	303
6.6	Presentation of findings relative to anthropometric reference data	305
6.6.1	For individuals	305
6.6.2	For populations	306
6.7	Recommendations	306
6.7.1	For Member States	306

6.7.2	For WHO	306
6.7.3	For future research and collection of reference data	307
	References	308
<b>7.</b>	<b>Overweight adults</b>	<b>312</b>
7.1	Introduction	312
7.1.1	Background	312
7.1.2	Biological and social significance of overweight	316
7.1.3	Anthropometry as an indicator of nutritional and health status	327
7.2	Using anthropometry in individuals	328
7.2.1	Screening for interventions	329
7.2.2	Assessing response to an intervention	330
7.3	Using anthropometry in populations	331
7.3.1	Targeting interventions	331
7.3.2	Assessing response to an intervention	331
7.3.3	Ascertaining determinants of overweight	332
7.3.4	Ascertaining consequences of overweight	332
7.3.5	Nutritional surveillance	333
7.4	Population data management and analysis	334
7.5	Potential development of reference data	336
7.6	Recommendations	339
7.6.1	For practical implementation	339
7.6.2	For future research	339
	References	340
<b>8.</b>	<b>Thin adults</b>	<b>345</b>
8.1	Introduction	345
8.1.1	Background	345
8.1.2	Terminology	345
8.2	Biological and social significance of anthropometry	346
8.2.1	Biological and social determinants of anthropometry	346
8.2.2	Biological and social consequences of anthropometry	347
8.3	Anthropometry as an indicator of nutritional and health status	349
8.3.1	Work capacity	349
8.3.2	Work productivity	349
8.3.3	Mortality at low body weight	351
8.3.4	Morbidity and low body weight	353
8.4	Interpretation of anthropometry	355
8.4.1	Considerations of body shape	355
8.4.2	Low body weight and body composition	356
8.5	Using anthropometry in individuals	359
8.6	Using anthropometry in populations	359
8.6.1	Targeting interventions	359
8.6.2	Assessing response to an intervention	359
8.6.3	Ascertaining determinants of malnutrition	360
8.6.4	Nutritional surveillance	360
8.6.5	Thinness as a public health problem	361
8.7	Guidelines for use of anthropometric indicators	362
8.7.1	Use of BMI with simple cut-off points	362
8.7.2	Arm and arm muscle circumference	364
8.7.3	Populations for which the guidelines may not be appropriate	368
8.8	Recommendations	369
8.8.1	For practical implementation	369
8.8.2	For future research	369
	References	370

9. Adults 60 years of age and older	375
9.1 Introduction	375
9.1.1 Background	375
9.1.2 Population variation in anthropometry	378
9.1.3 Anthropometry as an indicator of nutritional and health status	383
9.1.4 Interpretation issues in the elderly	389
9.2 Using anthropometry in individuals	390
9.2.1 Screening for interventions	390
9.2.2 Assessing response to an intervention	391
9.2.3 Assessment of functional ability	392
9.3 Using anthropometry in populations	393
9.3.1 Targeting interventions	393
9.3.2 Assessing response to an intervention	393
9.3.3 Ascertaining the determinants of thinness and overweight	394
9.3.4 Ascertaining the consequences of thinness and overweight	394
9.3.5 Nutritional surveillance	394
9.4 Methods of taking measurements	395
9.4.1 Weight	395
9.4.2 Height	396
9.4.3 Calf circumference	398
9.4.4 Subscapular skinfold thickness	398
9.4.5 Mid-upper arm circumference	398
9.4.6 Triceps skinfold thickness	399
9.5 Sources and characteristics of reference data	399
9.6 Recommendations	405
9.6.1 For practical implementation	405
9.6.2 For future research	405
References	407
10. Overall recommendations	410
10.1 For Member States	410
10.2 For WHO	410
10.3 For research	411
Acknowledgements	412
Annex 1	
Glossary of terms and abbreviations	416
Annex 2	
Recommended measurement protocols and derivation of indices	424
Annex 3	
Recommended reference data	439

Please click on the chapter headings for access to the full text.

## TABLE OF CONTENTS

1. Introduction.....	1
2. Technical Framework .....	4
<b><u>3. Pregnant and lactating women</u></b> .....	<b><u>37</u></b>
<b><u>4. The newborn infant</u></b> .....	<b><u>121</u></b>
<b><u>5. Infants and children</u></b> .....	<b><u>161</u></b>
<b><u>6. Adolescents</u></b> .....	<b><u>263</u></b>
<b><u>7. Overweight adults</u></b> .....	<b><u>312</u></b>
<b><u>8. Thin adults</u></b> .....	<b><u>345</u></b>
<b><u>9. Adults 60 years of age and older</u></b> .....	<b><u>375</u></b>
<b><u>10. Overall recommendations</u></b> .....	<b><u>410</u></b>
<b><u>Annexes</u></b> .....	<b><u>416</u></b>



# **WHO Expert Committee on Physical Status: The Use and Interpretation of Anthropometry**

Geneva, 1–8 November 1993

## **Members**

- Dr A. Ferro-Luzzi, Unit of Human Nutrition, National Institute of Nutrition, Rome, Italy
- Dr C. Garza, Director, Division of Nutritional Sciences, Cornell University, Ithaca, NY, USA
- Dr J. Haas, Division of Nutritional Sciences, Cornell University, Ithaca, NY, USA
- Dr J.-P. Habicht, Division of Nutritional Sciences, Cornell University, Ithaca, NY, USA  
(*Chairman*)
- Dr J. Himes, Division of Epidemiology, School of Public Health, University of Minnesota, Minneapolis, MN, USA (*Co-Rapporteur*)
- Dr A. Pradilla, Department of Epidemiology, University of Valle, Cali, Colombia
- Dr L. Raman, National Institute of Nutrition, Indian Council of Medical Research, Hyderabad, India
- Dr O. Ransome-Kuti, Former Professor of Paediatrics, University of Lagos, Lagos, Nigeria
- Dr J. C. Seidell, Head, Department of Chronic Diseases and Environmental Epidemiology, National Institute of Public Health and Environmental Protection, Bilthoven, Netherlands (*Co-Rapporteur*)
- Dr C. Victora, Department of Social Medicine, Faculty of Medicine, Federal University of Pelotas, Pelotas, Brazil (*Co-Rapporteur*)
- Dr M. L. Wahlqvist, Monash Medical Centre, Monash University, Melbourne, Australia
- Dr R. Yip, Chief, Maternal and Child Nutrition, Centers for Disease Control, Atlanta, GA, USA

## **Representatives of other organizations**

*Food and Agriculture Organization of the United Nations (FAO)*

Dr R. Weisell, Nutrition Officer, Food Policy and Nutrition Division, Rome, Italy

*United Nations Children's Fund (UNICEF)*

Dr J. Csete, Nutrition Section, New York, NY, USA

## **Secretariat**

- Dr G. A. Clugston, Chief Medical Officer, Nutrition, WHO, Geneva, Switzerland
- Dr M. de Onis, Scientist, Nutrition, WHO, Geneva, Switzerland (*Secretary*)
- Dr P. Eveleth, Former Deputy Associate Director, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA (*Temporary Adviser*)
- Dr M. Kramer, Department of Paediatrics and of Epidemiology and Biostatistics, McGill University, Montreal, Canada (*Temporary Adviser*)
- Dr P. Sizonenko, Head, Division of Biology of Growth and Reproduction, Cantonal University Hospital, Geneva, Switzerland (*Temporary Adviser*)

Dr J. Tuomilehto, Department of Epidemiology, National Public Health Institute,  
Helsinki, Finland (*Temporary Adviser*)

Dr J. Villar, Medical Officer, Special Programme of Research, Development and  
Research Training in Human Reproduction, WHO, Geneva, Switzerland

## Abbreviations

The following abbreviations are used in this report:

AGA	appropriate-for-gestational-age
AMA	arm muscle area
AMC	arm muscle circumference
API	adequate ponderal index
BMI	body mass index
EF	etiologiical fraction
IUGR	intrauterine growth retardation
LBW	low birth weight
LGA	large-for-gestational-age
LMP	last menstrual period
LPI	low ponderal index
MUAC	mid-upper arm circumference
NCHS	National Center for Health Statistics
NHANES	National Health and Nutrition Examination Survey
OR	odds ratio
PIH	pregnancy-induced hypertension
PPV	positive predictive value
ROC	receiver (or relative) operating characteristics
RR	relative risk
SD	standard deviation
SE	sensitivity
SF	symphysis-fundus (height)
SGA	small-for-gestational-age
SP	specificity
VLBW	very low birth weight

## 1. **Introduction**

The WHO Expert Committee on Physical Status: The Use and Interpretation of Anthropometry met in Geneva from 1 to 8 November 1993. Dr F.S. Antezana, Assistant Director-General, opened the meeting on behalf of the Director-General. This meeting was the culmination of a two-year preparatory process, involving more than 100 experts worldwide, which started in 1991 with the establishment of seven subcommittees, and continued with a number of subcommittee meetings, small group workshops, individual and working group contributions, and external reviews. During this process, subcommittees not only reviewed the latest knowledge in their areas, but also, in some cases, moved forward by themselves contributing to the knowledge pool.

The subcommittees received support for their efforts from numerous institutions, organizations, and governments. WHO takes pleasure in drawing attention to these contributions, without which many of the major preparatory activities would have been impossible. All individuals and institutions who contributed to the work are recorded in the Acknowledgements section (page 412).

Each year, 26 million babies are born too small to lead healthy lives, because their mothers were either ill or malnourished. More than 230 million (43%) of all preschool children in the developing world are stunted in their growth because of malnutrition caused by lack of food and by disease. Today, it is expected that this malnutrition will kill about seven million children a year, either directly or by worsening the impact of infectious diseases.

About 15% of non-elderly adults are too thin because of malnutrition and disease, which decrease their productivity and double their rate of premature mortality. At the same time, 150 million adults are overweight, of whom 15 million will die prematurely because of diseases resulting from obesity. In some communities almost all cases of adult diabetes and 40% of cases of coronary heart disease are attributable to body weight in excess of the optimum.

Data such as these on low birth weight, stunting, thinness, and overweight are obtained from measurements of height and weight. Anthropometric measurements assess body size and composition, and reflect inadequate or excess food intake, insufficient exercise, and disease. They demonstrate that deprivation and excess may coexist not only across, but also within, countries and even households, and show too that certain kinds of development and health policy enhance nutrition while others do not. Simple body measurements also permit the selection of individuals, families, and communities for interventions designed to improve not only nutrition but health in general and thus survival.

Anthropometry is the single most universally applicable, inexpensive, and non-invasive method available to assess the size, proportions, and composition of the human body. Moreover, since growth in children and

body dimensions at all ages reflect the overall health and welfare of individuals and populations, anthropometry may also be used to predict performance, health, and survival. This report describes appropriate uses and interpretation of anthropometry from infancy to old age. These applications are important for public health and clinical decisions that affect the health and social welfare of individuals and populations.

Over the years, WHO and other specialized agencies of the United Nations system have sought to provide guidance on the appropriate uses of anthropometric indices (1-6). Previously, attention has been focused largely on infants and young children, because of their vulnerability, and on the value of anthropometry in characterizing growth and well-being. Advances during the past decade, however, have demonstrated the relevance of anthropometry throughout life, not only for individual assessments but also for reflecting the health status and social and economic circumstances of population groups. In recognition of these developments, WHO convened an Expert Committee to re-evaluate the value of anthropometric indices and indicators at different ages in assessing health, nutrition, and social well-being. The Expert Committee recognized different needs and applications through the life cycle, and addressed these issues as they relate to pregnant and lactating women, the newborn, infants and children, adolescents, adults, and elderly people (aged 60 years or more).

Paediatricians have long used child growth as an important parameter in evaluating the health and general well-being of children (7). In the nutrition field, low height and/or weight relative to reference data have been used as classic indicators of undernutrition for individuals and groups; similarly, elevated body weight and thickness of subcutaneous fat have become common indicators of overnutrition or obesity.

Recent research has expanded the applications of anthropometry to include predicting who will benefit from interventions, identifying social and economic inequity, and evaluating responses to interventions. Importantly, it has become clear that different uses of anthropometry require different properties of the most appropriate anthropometric indicators, and that appropriate applications and interpretations of anthropometric indicators may be different for individuals and for populations. Further, appropriate indicators for a particular purpose may vary according to the prevalence of a specific problem.

Principles of public health screening (8) and epidemiology are particularly helpful in identifying appropriate anthropometric indicators, and specifying optimum cut-off points for variables (9). Experience with surveillance (2) has contributed to concepts and practices concerning community assessments and “trigger-levels” as a basis for public health decisions.

The Expert Committee was requested to:

- develop recommendations for the appropriate use and interpretation of anthropometry in individuals and populations in various operational settings;

- identify and/or develop reference data for anthropometric indicators when appropriate;
- provide guidelines on how these reference data should be used; and
- identify new or unresolved issues and gaps in knowledge that require further research.

The Expert Committee's report is intended to provide a framework and contexts for present and future uses and interpretation of anthropometry. Technical aspects of this framework are presented in section 2, and specific applications of anthropometry appropriate for a particular physical status or for particular age groups are dealt with in subsequent sections. For some groups, such as adolescents and the elderly, there has been little previous research, and the report provides a basis and impetus for future studies. For other age groups, such as infants and children, the report provides a re-evaluation in the light of current research, and allows for an integrated approach to anthropometry throughout life. It is intended to furnish scientists, clinicians, and public health professionals worldwide with an authoritative review, reference data, and recommendations for the use and interpretation of anthropometry that should be appropriate in many settings.

## References

1. Jelliffe DB. *The assessment of the nutritional status of the community*. Geneva, World Health Organization, 1966 (WHO Monograph Series, No. 53).
2. *Methodology of nutritional surveillance. Twenty-seventh Report of a Joint FAO/UNICEF/WHO Expert Committee*. Geneva, World Health Organization, 1976 (WHO Technical Report Series, No. 593).
3. Waterlow JC et al. The presentation and use of height and weight data for comparing nutritional status of groups of children under the age of 10 years. *Bulletin of the World Health Organization*, 1977, 55:489-498.
4. *Measuring change in nutritional status: guidelines for assessing the nutritional impact of supplementary feeding programmes*. Geneva, World Health Organization, 1983.
5. WHO Working Group. Use and interpretation of anthropometric indicators of nutritional status. *Bulletin of the World Health Organization*, 1986, 64(6): 929-941.
6. Beaton G 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).
7. Tanner JM. *History of the study of human growth*. Cambridge, Cambridge University Press, 1981.
8. Wilson JMG, Jungner G. *Principles and practice of screening for disease*. Geneva, World Health Organization, 1968 (Public Health Papers, No. 34).
9. Galen RS, Gambino SR. *Beyond normality: the predictive value and efficiency of medical diagnoses*. New York, Wiley, 1975.

## 2. **Technical framework**

### 2.1 **Introduction**

Anthropometry has been widely and successfully applied to the assessment of health and nutritional risk, especially in children. Recent publications have refined the interpretation of anthropometric indicators in selected operational settings (1), but little guidance has been published concerning other appropriate uses of anthropometry. The implications of specific uses for the choice of indicators and interpretation of findings are not fully understood, even though correct selection of the best anthropometric indicators depends entirely on the purposes for which they are used (2).

This section deals with the technical basis underlying the various uses of anthropometric indicators, using principles of applied biostatistics and epidemiology. For the broader audience, these principles are explained without equations; readers interested in a technically more sophisticated treatment are referred to the specialized readings cited.

### 2.2 **Levels of body composition**

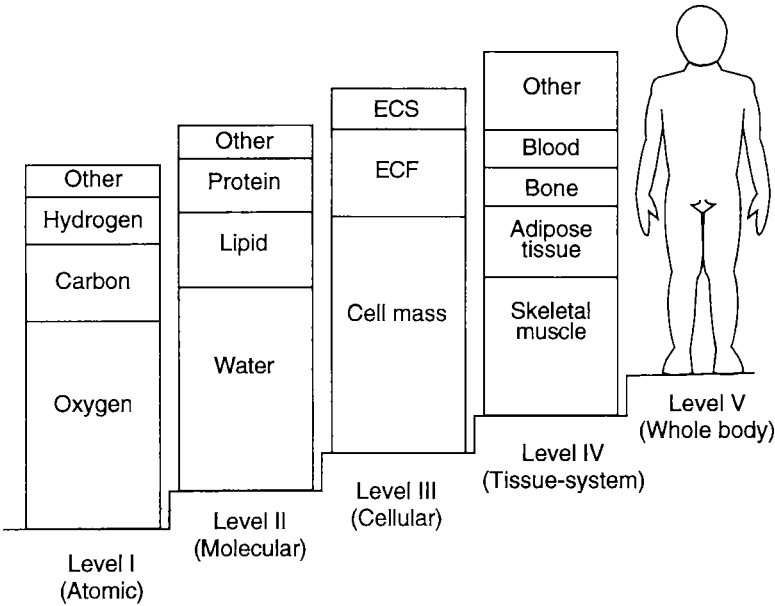
Full appreciation of the utility of anthropometry requires an understanding of the organizational levels of human body composition. Recently, there have been major advances in conceptual models relating anthropometry to body composition, which provide insight into the physiological mechanisms represented by anthropometry (3).

The five organizational levels of body composition and their major compartments are shown in Fig. 1. At the atomic level, the major chemical elements are oxygen, hydrogen, carbon, nitrogen, calcium, and phosphorus. Whole-body measurements of these constituents are usually made with research techniques such as neutron activation analysis, and provide important information. For example, nitrogen balance is an indicator of protein turnover, and total body calcium is an indicator of total bone mineral.

The next level of body composition comprises the major molecular compartments such as water, protein, glycogen, mineral (osseous and non-osseous), and fat (Fig. 2). Water and osseous minerals can be measured directly, but fat, protein, glycogen, and non-osseous minerals must be estimated by indirect techniques. Each of the several methods used to estimate this latter group of constituents relies on assumptions that relate measurable aspects of body composition to the constituent of interest. Anthropometric methods of estimating total body fat and fat-free mass (FFM) are usually developed using one of these indirect techniques.

The cellular level of body composition consists of cells, extracellular fluid (ECF), and extracellular solids (ECS). A widely used model

Figure 1  
**Five-level model of body composition<sup>a</sup>**



Note: ECS = extracellular solids  
 ECF = extracellular fluid

<sup>a</sup> Reproduced from reference 3 with the permission of the American Society for Clinical Nutrition.

considers the total cellular mass to be composed of two components – fat (a molecular-level compartment), and the fat-free cell mass referred to as body cell mass (BCM), where most metabolic processes take place. Cells are the body’s main functional compartments. Several equations based on anthropometry have been developed to predict body cell mass at the cellular level, although their accuracy is a matter of debate and none is widely used.

The tissue-system level of body composition consists of the major tissues, organs, and systems; thus body weight is equal to adipose tissue + skeletal muscle + bone + blood + residual (visceral organs, etc.). Adipose tissue includes adipocytes, blood vessels, and structural elements, and is the primary site of lipid storage. It is located mainly in the subcutaneous and internal or visceral compartments, with its distribution under hormonal and genetic control.

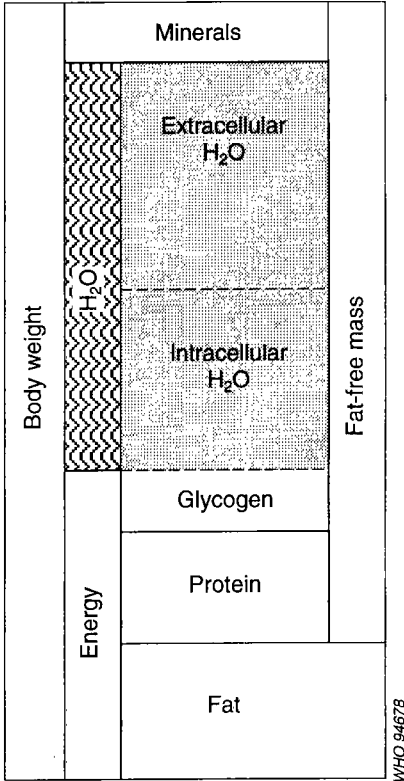
A steady-state relationship exists between the various body-composition compartments. That is, there are stable quantitative relationships between compartments at the same and different levels of body composition that remain relatively constant over a specified time (usually months or years). This permits information about body composition at various



Figure 2

**The major components of body weight<sup>a</sup>**

Water, protein, and mineral within the fat-free body mass occur in the average proportions 0.725, 0.195, and 0.08; glycogen is variable at 0.01 to 0.02; 50 to 55% of water is intracellular, with the remainder in the extracellular space.



<sup>a</sup>Adapted from reference 3 with the permission of the American Society for Clinical Nutrition.

levels to be derived from anthropometric measurements made at the whole-body level. Both aging and disease affect these quantitative relationships, and anthropometry provides a means of detecting the resultant changes.

**2.3 Anthropometric measurements, indices, and indicators**

**2.3.1 Measurements**

The basic anthropometry measurements considered here are weight and height, but principles derived from these measures may be applied to other measurements. The methods for collecting recommended data are presented in Annex 2.

### 2.3.2 Indices

Anthropometric indices are combinations of measurements. They are essential for the interpretation of measurements: it is evident that a value for body weight alone has no meaning unless it is related to an individual's age or height (4). Thus, for example, measurements of weight and height may be combined to produce the body mass index (weight/height<sup>2</sup>) or a ponderal index (weight/height<sup>3</sup>), or weight may be related to height through the use of reference data. In children, the three most commonly used anthropometric indices are weight-for-height, height-for-age, and weight-for-age; other indices are used for different age/physiological groups, such as pregnancy weight gain in pregnant women.

The anthropometric indices can be expressed in terms of Z-scores, percentiles, or percent of median, which can then be used to compare a child or group of children with a reference population. These reporting systems are defined as follows:

- Z-score (or standard deviation score) (5, 6) – the deviation of the value for an individual from the median value of the reference population, divided by the standard deviation for the reference population:

$$\text{Z-score or SD-score} = \frac{(\text{observed value}) - (\text{median reference value})}{\text{standard deviation of reference population}}$$

A fixed Z-score interval implies a fixed height or weight difference for children of a given age. A major advantage of this system is that, for population-based applications, it allows the mean and standard deviation to be calculated for a group of Z-scores.

- Percentile – the rank position of an individual on a given reference distribution, stated in terms of what percentage of the group the individual equals or exceeds. Thus a child of a given age whose weight falls in the 10th percentile weighs the same or more than 10% of the reference population of children of the same age.

Percentiles are commonly used in clinical settings because their interpretation is straightforward. However, the same interval of percentile values corresponds to different changes in absolute height or weight, according to which part of the distribution is concerned, and it is therefore inappropriate to calculate summary statistics such as means and standard deviations for percentiles. Moreover, towards the extremes of the reference distribution there is little change in percentile values, when there is in fact substantial change in weight or height status.

- Percent of median – the ratio of a measured value in the individual, for instance weight, to the median value of the reference data for the same age or height, expressed as a percentage.

The main disadvantage of this system is the lack of exact correspondence with a fixed point of the distribution across age or height status. For example, depending on the child's age, 80% of the median weight-for-age might be above or below  $-2$  Z-scores; in terms of health, this would result in different classification of risk. In addition, typical cut-offs for percent of median are different for the different anthropometric indices; to approximate a cut-off of  $-2$  Z-scores, the usual cut-off for low height-for-age is 90%, and for low weight-for-height and low weight-for-age 80%, of the median (7-9).

If the distribution of reference values follows a normal (bell-shaped or Gaussian) distribution, percentiles and Z-scores are related through a mathematical transformation. The commonly used  $-3$ ,  $-2$ , and  $-1$  Z-scores are, respectively the 0.13th, 2.28th, and 15.8th percentiles. Similarly, the 1st, 3rd and 10th percentiles correspond to, respectively, the  $-2.33$ ,  $-1.88$ , and  $-1.29$  Z-scores. It can be seen that the 3rd percentile and the  $-2$  Z-score are very close to each other.

The main characteristics of the three reporting systems are summarized and compared in Table 1. A detailed treatment of their limitations and strengths is to be found elsewhere (9); the brief discussion above is intended to outline the reasons for the Z-score being the preferred system.

The use of indices derived from reference data is appropriate for many purposes, but for other purposes there are better ways of adjusting anthropometric values for age and sex, such as through multivariate analysis (5) or residual analysis (10). However, these methods are in general more suitable for research applications and will not be further discussed here.

It is important to note that all indices derived from age-specific reference data depend for their precision on exact knowledge of age; when this information is not available, use of age-based indices such as height-for-age may result in misclassification (11).

### 2.3.3 *Indicators*

The term "indicator" relates to the use or application of indices. The indicator is often constructed from indices; thus, the proportion of children below a certain level of weight-for-age is widely used as an indicator of community status.

The anthropometric indices discussed here all relate to body size and composition. Sometimes this is the only type of relationship that can be inferred; indices should then be referred to as body size or body composition indicators, rather than as nutrition or health indicators. Depending on the circumstances, the same anthropometric index may be influenced equally by nutrition and health, or more by one than by the

Table 1

**Comparison of the characteristics of three anthropometric data-reporting systems**

Characteristic	Z-score	Percentile	Percent of median
Adherence to reference distribution	Yes	Yes	No
Linear scale permitting summary statistics	Yes	No	Yes
Uniform criteria across indices	Yes	Yes	No
Useful for detecting changes at extremes of the distributions	Yes	No	Yes

other; accordingly it may then be referred to as an indicator of nutrition, or of health, or of both. In some cases, the index may be used as a distal, or indirect, indicator of socioeconomic status or of inequities in socioeconomic status; if the index is genuinely influenced by these factors, even though indirectly through nutrition and health, it may then be referred to as a socioeconomic or equity indicator.

A valid nutritional indicator owes a substantial proportion of its variability to differences in nutrition. For any given indicator, however, this proportion may vary across or within populations. For instance, body mass index (BMI), the ratio of weight to the square of height, is a good indicator of variability in energy reserves in individuals with a sedentary lifestyle, but not in athletes; similarly, low birth weight reflects maternal malnutrition in mothers who are too thin, but not in mothers who are overweight.

It is not uncommon for an indicator to be erroneously interpreted as reflecting nutrition or some other factor, when this is not the case. This may lead to inappropriate targeting of intervention programmes. For example, providing energy supplements to mothers in a particular area on the basis of the prevalence of low birth weight alone will not succeed if smoking is common in the area. For low birth weight to be useful as an indicator of nutritional status within this population, it must be “conditioned” on the nutritional status of the mothers. That is, other factors must be taken into account in assessing the nutritional status of populations from indicators thought to be nutritional. Thus, the prevalence in the population of the nutritional or health factor of concern *conditions* the interpretation of an anthropometric indicator.

Choice and conditioning of indicators should ultimately depend on the decisions that will be made on the basis of the information they yield. Throughout, this report attempts to relate the indicators to the actions that will be taken on behalf of individuals or populations.

## 2.4 Selection of anthropometric indicators

Anthropometric indicators can be classified according to the objectives of their use, which include the following (the order of listing is dictated by various methodological considerations discussed later):

- *Identification of individuals or populations at risk.* In general, this requires data based on indicators of impaired performance, health, or survival. Depending on the specific objective, the anthropometric indicators must:
  - reflect past or present risk, or
  - predict future risk.

An indicator may reflect both present and future risk; for instance, an indicator of present malnutrition may also be a predictor of an increased risk of mortality in the future. However, a reflective indicator of past problems may have no value as a predictor of future risk; for example, stunting of growth in early childhood as a result of malnutrition may persist throughout life (1), but with age probably becomes less reliably predictive of future risk.

Indicators of this type might be used in the risk approach to identification of health problems and potential interventions (12), although, as discussed below, the risk approach may have little value in predicting the *benefit* to be derived from interventions. An indicator of risk could, however, be appropriately used to assign higher life-insurance rates to obese individuals because of their increased risk of death.

- *Selection of individuals or populations for an intervention.* In this application, indicators must:
  - predict the benefit to be derived from the intervention.

The distinction between indicators of risk and indicators of benefit is not widely appreciated, yet it is paramount for developing and targeting interventions. Some indicators of present or future risk may also predict benefit, but this is not necessarily the case. Low maternal height, for example, predicts low birth weight, but, in contrast to low maternal weight in the same population, does not predict any benefit of providing an improved diet to pregnant women. By the same token, predictors of benefit may not be good predictors of risk.

Anthropometry provides important indicators of overall socioeconomic development among the poorest members of a population. Stunting in children and adults reflects socioeconomic conditions that are not conducive to good health and nutrition: thus stunting in young children may be used effectively to target development programmes.

- *Evaluation of the effects of changing nutritional, health, or socioeconomic influences, including interventions.* For this purpose indicators must:
  - reflect response to past and present interventions.

Change of weight-for-height is a good example of an indicator of response in a wasted child being treated for malnutrition. At the population level a decrease in the prevalence of stunting is an indicator that social development is benefiting the poor as well as the comparatively affluent. Similarly, a decrease in the prevalence of low birth weight would indicate success in controlling malaria during pregnancy (13).

In describing an indicator of response, the possible lag between the start of an intervention and the time when a response becomes apparent is an important consideration. At the individual level, a wasted infant will respond to improved nutrition first by putting on weight and then by “catching up” in linear growth. At the population level, however, decades may elapse before improvements can be seen in adult height (14).

- *Excluding individuals from high-risk treatments, from employment, or from certain benefits.* Decisions regarding an individual’s inclusion in, or exclusion from, a high-risk treatment protocol, consideration for employment in a particular setting (e.g. an occupation requiring appreciable physical strength), or admission to certain benefits (e.g. low life-insurance rates) depend on indicators that:

- predict a lack of risk.

Anthropometric indicators of lack of risk were once presumed to be the same as those that predict risk, but recent work has revealed that this is not invariably the case (15). In the cited studies, indicators of poor growth were less effective in predicting adequate growth than other indicators.

- *Achieving normative standards.* Assessing achievement of normative standards requires indicators that:

- reflect “normality”.

Some activities appear to have no objectives beyond encouraging individuals to attain some norm. For instance, moderate obesity among the elderly is not associated with poor health or increased risk of mortality, and weight control in this age group is therefore based solely on normative distributions.

- *Research purposes that do not involve decisions affecting nutrition, health, or well-being.* The indicator requirements for these objectives, whether they concern individuals or whole populations, are generally beyond the scope of this report. The need to build appropriate biological, behavioural, and epidemiological models into the analyses often means that some simpler indicators, including some discussed in this report, may be inadequate for research purposes.

There may be differences in the interpretation of anthropometric indicators when applied to individuals or to populations. For example,

while a reflective indicator, such as the presence of marasmus, signifies malnutrition in a given child today, a sudden increase of marasmus in a population may be predictive of future famine. The appropriateness of indicators thus depends on the specific objectives of their use, and research is only just beginning to address this specificity and its implications. Little is known, for example, about how the use of different cut-offs for anthropometric indicators fulfils different objectives. Consequently, this report must be largely tentative in its recommendations concerning the coupling of indicators and applications, and should be regarded as a basis for future improvements in research.

## 2.5 Sensitivity and specificity of indicators

A good indicator is one that best reflects the issue of concern or predicts a particular outcome. Discussion of the methodology for choosing appropriate indicators and cut-offs focuses on risks to health because little work has been done on selection of indicators for other objectives. Risk of mortality is used in the following examples.

Historically, two approaches have been used to identify anthropometric variables that show an association with death. The first uses classical statistical methods that describe relationships between anthropometric indices and death (e.g. ordinary least squares, and logistic regression (5)). The second, discussed below, is based simply on how well the indicator separates those who will die from those who will survive, and is intuitively more obvious to practitioners in public health who are concerned with screening. The two approaches are, however, related (16).

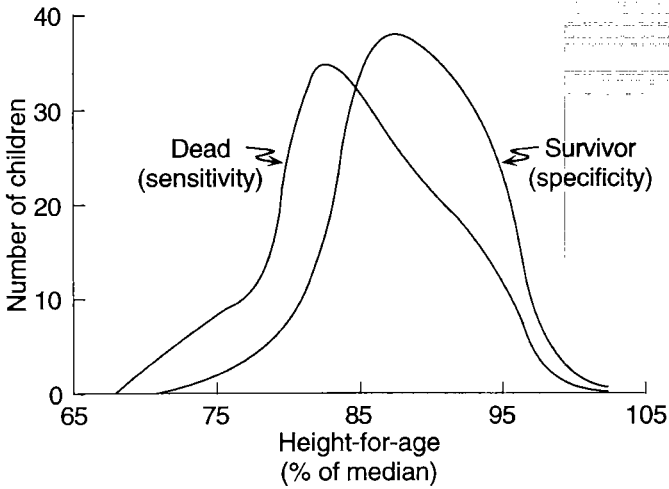
A screening test identifies individuals at risk on the basis of an indicator and a specific cut-off point. Of those who will die, the proportion who are identified as cases by the test is a measure of the *sensitivity* of the screening test. Sensitivity can be improved by changing the cut-off point to identify more people as being at risk. This is illustrated in Fig. 3, which shows the numbers of children at different height-for-age values who died – that is, the sensitivity frequency distribution. As the value of the cut-off point is raised from 65% of median to 100% of median, the sensitivity increases from 0 to 100%, because more children are diagnosed as being at risk. This sensitivity distribution is presented cumulatively in Fig. 4.

Because the sensitivity of a test changes with the cut-off point, sensitivity alone cannot be used in comparing indicators. It is also essential to consider the performance of the screening test in accurately excluding those who will not die. This is the *specificity* of the test; its frequency distribution is also shown in Fig. 3. As the cut-off point is lowered from 105% of median to 67% of median, the proportion of individuals excluded by the test rises, as shown in Fig. 4.

A considerable overlap in anthropometric values between the sensitivity and specificity distributions is apparent in Fig. 3. This is to be expected,

Figure 3

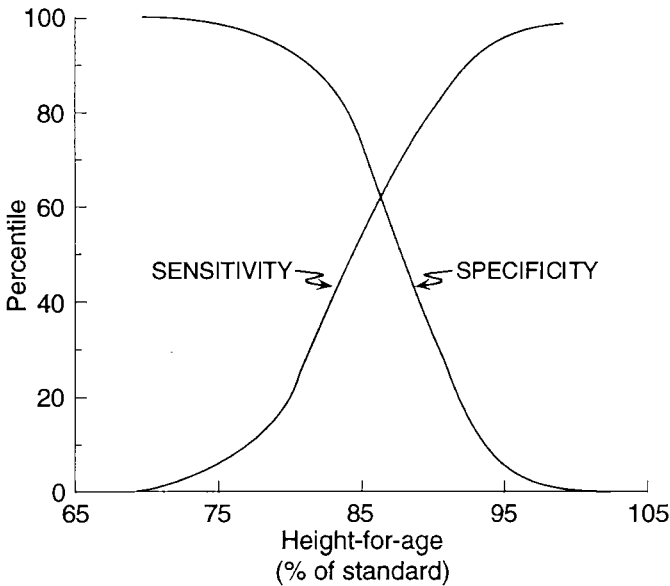
**Percentage frequency distribution of height-for-age according to 2-year survival<sup>a</sup>**



<sup>a</sup> Adapted from reference 17 with the permission of the American Society for Clinical Nutrition.

Figure 4

**Percentile plot of sensitivity and specificity values for height-for-age as calculated from Fig. 3<sup>a</sup>**



<sup>a</sup> Adapted from reference 17 with the permission of the American Society for Clinical Nutrition.



since sensitivity and specificity are inversely related: increasing one (by changing the cut-off point) results in a decrease in the other, as seen in Fig. 4.

Values for sensitivity and specificity are often assumed to be constant for indicators, unaffected by the prevalence or incidence of the condition of interest. Other descriptors of screening tests, however, such as the positive predictive value, are affected by prevalence, which makes them inappropriate for use in comparing indicators across different populations (18).

Specificity and sensitivity will be affected by the underlying biological and behavioural processes that relate the indicator to the outcomes of interest in different settings. For example, in a setting where low birth weight is due mainly to prematurity – which is strongly associated with the early neonatal death rate – its sensitivity as an indicator of mortality will be greater than where it is due mainly to intrauterine growth retardation, which is less strongly associated with the death rate. Thus, for outcomes that may be influenced by several factors, variability in the sensitivity and specificity distributions is to be expected (19). More consistency may be expected of indicators that predict or measure response to an intervention with a well established outcome. For this reason, changes in weight and mid-upper arm circumference are more sensitive than height to short-term seasonal influences (20), but height is generally more responsive than weight to improved food intake in the long term (21).

## 2.6 Selection of a best indicator

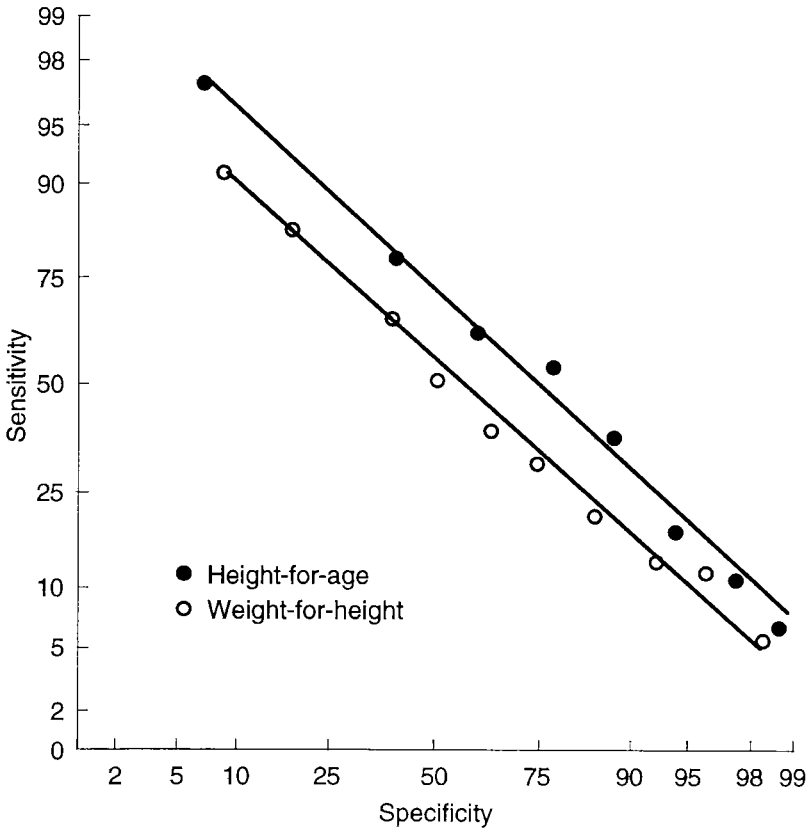
The trade-off between sensitivity and specificity can be represented graphically by plotting probability values for sensitivity against those for specificity at various cut-off points (see Fig. 5) to produce a curve of “receiver (or relative) operating characteristics” (ROC) that permits a comparison of indicators over their whole range. The curves have been linearized by the Z-transformation (5).

For a given specificity, height-for-age has greater sensitivity than weight-for-height (Fig. 5) in identifying those who will die in the subsequent 2 years within the population studied. Accordingly, height-for-age generates fewer errors of classification at every level of sensitivity and specificity than weight-for-height. In this context, therefore, height-for-age is the better indicator. If the curves in the Z-transformed ROC presentation are parallel, one indicator is obviously best over all ranges; if the ROC curves cross, however, one indicator is probably better over one range of values, while the other is better over another range.

Statistical methods for selecting a best indicator have been outlined elsewhere (16). When indicators are to be compared in an effort to select

Figure 5

**Relationship between sensitivity (%) and specificity (%) for height-for-age and weight-for-height in predicting 2-year survival<sup>a</sup>**



<sup>a</sup> Reproduced from reference 17 with the permission of the American Society for Clinical Nutrition.

the best, results can be misleading if the cut-off point is chosen *before* selection of a particular indicator: there is no *a priori* best cut-off for purposes of comparison.

In the example above, death and survival were the basis for dividing the population into sensitivity and specificity distributions respectively. For identifying the best indicator of response, the population should be divided on the basis of whether or not the treatment was administered (22); to predict benefit, and to identify the characteristics associated with best response, the population should be divided into those who respond to the intervention and those who do not.

## 2.7 Using anthropometry in individuals

At the level of the individual, anthropometry is used either to identify a person as being in need of special consideration, or to assess that person's response to some intervention.

### 2.7.1 Screening with one measurement for targeting an intervention

Screening for malnutrition or disease is of value only if effective treatment for the condition is available (23). This principle is frequently overlooked in the context of anthropometric screening for malnutrition. Moreover, the anthropometric screen may be the *only* step taken before decisions on intervention are reached, particularly in emergencies. Historically, however, screening has been viewed as simply a *first* step (23, 24), i.e. as the first in a sequence of increasingly specific screens leading to effective intervention. It is therefore clear that, the less specific the initial screens are for the intervention envisaged, the more important the subsequent screens become.

An anthropometric screen is based on an indicator for which a suitable cut-off point (or points) is chosen to categorize individuals for different decisions. The crucial questions to be answered by an anthropometric screen are:

- Is the indicator the best one for the decision that must be made? and
- Are the cut-off points the best ones for selecting individuals and ensuring the necessary action?

The first question was addressed in section 2.6; the following section deals with selection of cut-off points.

#### *Selecting the best cut-off point*

Universal cut-off points are often recommended but are appropriate only if resources are adequate to handle all individuals selected for intervention and if the intervention causes no adverse side-effects. In such a case, it is unimportant that a high proportion of those who receive the intervention will not benefit from it. The cut-off point should be set at 100% sensitivity, so that all those at risk who *can* benefit from the intervention are treated.

Cut-offs are commonly set on the basis of experience in affluent populations which shows that the proportion of individuals identified by a screen who can benefit is sufficient to warrant further diagnostic steps. These cut-offs are usually described in terms of Z-scores, percentiles, or percent of a normative median because, historically, reference data from healthy populations were used to establish these values. However, these reference data give information only about healthy individuals who cannot benefit from the intervention; they provide no indication of the sensitivity that is relevant for setting the cut-off. Nonetheless, despite the

poor theoretical basis for using reference data in this way, these cut-offs have been tested empirically in affluent populations and are now conventional; they should not be abandoned until cut-offs based on sounder principles have been validated.

In certain situations, the specificity distribution of those who cannot benefit from an intervention is important: when the proportion of the population who could benefit is low and the intervention causes adverse side-effects. The introduction of supplementary feeding for a fully breast-fed infant, especially in areas where food contamination and infection are common, probably falls into this category, since the new foods may pose unnecessary health risks to the child (25, 26). It may be that some deficit in nutrition is less harmful than the introduction of potentially contaminated foods, in which case specificity would be more important than sensitivity in setting the cut-off.

#### *Balancing needs and resources in a population*

Where resources are insufficient to support intervention for *all* those who might need it, cut-offs should be chosen to maximize the number of at-risk individuals who *can* be treated. Ideally, screening for risk of death should maximize the coverage of those at risk, and at the same time maximize the proportion of those selected who are in fact at risk. The coverage is directly reflected by the sensitivity of the screen. The term “yield” (23) was formerly used for the proportion of those at risk among the total selected, but this has been superseded by “positive predictive value” (24, 27).

Sensitivity and positive predictive value can be readily confused, since the numerator of each ratio (the number of individuals who are at risk and who are identified by the screening test) is the same. This is illustrated in Table 2, where those who are at risk are represented by  $B$  and  $b$ , and those not at risk by  $N$  and  $n$ . Although the denominators of both proportions include the  $B$  individuals, they differ in the remaining individuals they include. For sensitivity, a proportion ( $b$ ) of the at-risk individuals are included, who should have been picked up by the screen and were not; they are false-negatives, i.e. individuals falsely diagnosed as not at risk. For the positive predictive value, a proportion ( $n$ ) of the individuals not at risk is included in the denominator; they are false-positives, i.e. individuals falsely diagnosed by the screen as being at risk.

As stated above, the most important principle in screening for an intervention where resources are limited is to select for treatment the greatest number of those who most need it (17). Thus, the screen should identify those in greatest need, and also have the highest positive predictive value. Such a screen will also capture the lowest proportion of individuals who do not need intervention. As presently conceived in anthropometry, the highest positive predictive value is best achieved by moving the cut-off as far as possible from the reference median value. Using risk of malnutrition as an example, the cut-off point for height-for-

Table 2

**Sensitivity, specificity, positive and negative predictive values, and prevalence**

		Truth		
		Will benefit	Will not benefit	Sum
Diagnosis based on screen	Will benefit	$B$	$n$	$B + n$
	Will not benefit	$b$	$N$	$b + N$
	Sum	$B + b$	$n + N$	$B + b + n + N$

Sensitivity =  $B/(B + b)$

True-positive =  $B$

Specificity =  $N/(n + N)$

True-negative =  $N$

Prevalence =  $(B + b)/(B + b + n + N)$

False negative =  $b$

Positive predictive value =  $B/(B + n)$

False positive =  $n$

Negative predictive value =  $N/(b + N)$

age can be set so low that everybody chosen is malnourished. However, there are many others in the population who are malnourished and who would be identified by a higher cut-off. Unfortunately, the higher cut-off also selects individuals who are small but not malnourished, and therefore has a lower positive predictive value. Thus there is usually a trade-off between the positive predictive value and the sensitivity. In many circumstances, when selection for malnutrition is based on low height and weight indices, each increment by which the cut-off is moved towards the reference median delivers the next best positive predictive value compared with the previous cut-off, and also selects those individuals who are the next most malnourished. This is why the cut-off that selects exactly the number of people that can be handled by the intervention selects the most malnourished, with the best positive predictive value, and is therefore the best that can be chosen where resources are limited.

For lack of an appropriate example in anthropometry, a clinical study has been chosen to illustrate this point. The objective was to determine how well early signs and symptoms among young children with diarrhoea predicted the risk of subsequent life-threatening dehydration (28). The sensitivity of each indicator (e.g. vomiting, fever, thirst, six or more stools/day) was plotted against the proportion of all children in the study population who would be selected for close follow-up and intensive intervention. If resources were available for intensive intervention in 25% of the children with diarrhoea, the indicator giving the highest sensitivity for that proportion would be selected. Of course the positive predictive values will change for all of these cut-offs as the prevalence of life-threatening diarrhoea changes, but the order of sensitivity of the indicators will be unchanged.

*Effect of prevalence on positive predictive value and implications for setting cut-off points*

In contrast to sensitivity and specificity, the positive predictive value (PPV) always depends on the prevalence of the issue of concern. The higher the prevalence, the higher the PPV for a given cut-off.

The PPV is determined by the sensitivity and the specificity of the test and by the prevalence of the condition of interest in the population being tested. As is evident from Table 2, the more specific a test, the better its PPV (and thus the greater the confidence that an individual with a positive test result has the condition of interest). The mathematical formula relating PPV to sensitivity (SE), specificity (SP) and prevalence (P) is based on Bayes' theorem of conditional probability (29).

$$\text{PPV} = \frac{(\text{SE}) (P)}{(\text{SE}) (P) + (1-\text{SP}) (1-P)}$$

This formula shows that a screening test performs well, with moderately high specificity (90%), if prevalence of the condition in the population tested is relatively high. At lower prevalences, however, the PPV drops to nearly zero for the same specificity, and the test is virtually useless for screening purposes. In summary, the interpretation of a positive test result varies from setting to setting, according to the estimated prevalence of the disease or condition of interest.

In a population where malnutrition among children greatly exceeds 50%, a wasted child is almost certainly malnourished and further diagnostic screening is superfluous. However, more practical information may be required before a specific intervention can be implemented. For instance, if the underlying problem is one of maternal knowledge rather than the unreliability of the food supply, the nature of the intervention will differ.

Diagnostic screening for risk and etiological screening thus have different (though related) objectives, but the primary determinant of appropriateness in both cases is the positive predictive value.

The basic requirement for using the same percentile or Z-score as the cut-off point is that it should have the same meaning in different individuals. This can be true only if the positive predictive value of the test is constant, which is rare. It must therefore be concluded that universal cut-off points are less useful than those based on the principle of selecting for intervention as many people as the available resources can handle. Thus, where resources are insufficient to treat everybody who is in need, which is the usual situation among both affluent and poor societies worldwide, there is no universally ideal cut-off.

Simple, practicable methodologies for choosing cut-offs that take account of local availability of resources as well as of the number of people who need the intervention have yet to be developed. Any such

methodology should also provide guidance about additional information that may be needed, the nature of which will vary with prevalence, risk, or benefits sought.

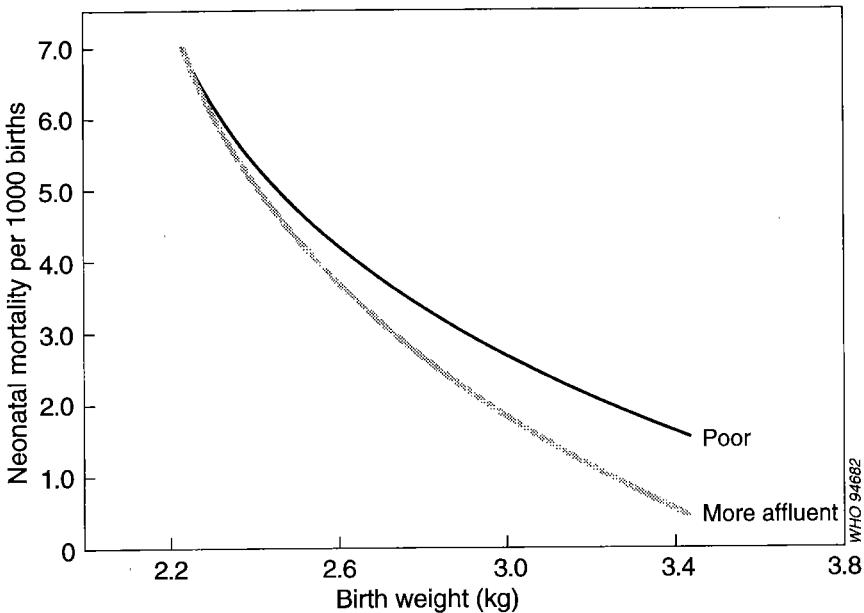
*Comparing cut-offs for identifying risk with those for predicting benefit*

Little research has examined the relationship between the most efficient cut-offs used to screen for risk and those used to predict benefit, but it seems likely that there are many circumstances in which the two are entirely unrelated. In Fig. 6, for example, neonatal mortality rates at different birth weights among the poor and a more affluent population of the same country are compared. If risk were the only guiding criterion, the most sensible intervention would target infants below, say, 2.5 kg, whether from affluent or poor families. However, assuming that the most efficient intervention would reduce mortality rates among the poor to the levels among the affluent, consideration of potential benefit will direct the intervention to those above this cut-off weight (30).

Thus, the choice and use of a cut-off to select individuals for special consideration may differ radically, depending on whether selection is based on prediction of death (risk indicator) or prediction of benefit from improved services (indicator of predicted benefit). This suggests again that the conventional approach of selecting on risk (12) and then following up with more specific screens (23) may generate conflicting decisions.

Figure 6

**Birth-weight-specific mortality of a poor and a more affluent population in the same country**



### *Growth monitoring and screening: changes in size over time*

Continuous monitoring is often undertaken for the early detection of health and nutrition problems, particularly during periods when the risk of malnutrition, morbidity, and death is high, as in early childhood or in old age. Measures of satisfactory progress are healthy growth in children and healthy maintenance of body mass in the elderly. In theory, both should be more easily verified by repeated measurements than by comparing attained size with the reference data.

This approach has been widely implemented among children within the larger context of growth monitoring. It is usual practice to ascertain whether the child is growing along a set percentile of reference data. At present, these reference data are cross-sectional, which poses certain problems, particularly in infancy and adolescence (31); further research is needed to develop appropriate longitudinal reference data.

Drawing up such data for growth increments is a formidable task, because the distribution of increments around the mean increment of the reference data will depend on the exact intervals between measurements. In practice, few children are measured at set intervals. In the future, it may be possible to compare individual growth data with figures from inexpensive computerized programs based on algorithms derived from reference data.

Even less research has focused on the relationship of growth faltering to responses to interventions than on small attained size, so that few quantitative conclusions can be drawn about the sensitivity distributions and positive predictive values essential for decisions on intervention. The receiver operating characteristics may be much better for incremental than for attained data: increments are a better reflection of present remediable circumstances than is attained size, and may be subject to less genetic variability. All these presumptions, of course, require empirical verification.

In establishing and maintaining a growth monitoring system, a number of other considerations may be just as important as the diagnosis of growth faltering itself. These include increased attention to child health and improvements in access to other services, in social networks, and in the early detection of diseases unrelated to growth.

### **2.7.2 Assessing response to an intervention**

In clinical practice change can be assessed from two or more serial measurements in the individual. Public health practice, by contrast, deals with populations and it is thus difficult reliably to assess change over time at the individual level.

Change may also be verified on the basis of a child's achieving some threshold. For instance, a wasted child may be selected to participate in a feeding programme. When the child has regained the level of a given



cut-off point – perhaps the same cut-off used in screening for the intervention – he or she can be discharged from the programme. This method, whereby the same screen is used to select for the intervention and to judge satisfactory response, usually shows a high rate of apparent response. However, if individuals are selected for the second measurement on the basis of the first, measuring the same individual twice will often result in spurious improvement. The greater the deviation of the first measurement from the population mean, the more likely it is that the individual will move towards that mean for reasons other than the intervention, including measurement errors and week-to-week variability. Careful account must be taken of this regression towards the mean (32). In one study in which they were considered in the context of a feeding programme, these other factors accounted for most of the response, even though prevalence of malnutrition was moderately high (33). While this does not detract from the value of the method as a basis for discharging children from a feeding programme, it must be taken into account in judging the effectiveness of the programme.

Anthropometry may also be used for deciding to discontinue an intervention in individuals who fail to respond. In such cases, medical examination may disclose other treatable causes of poor growth.

## **2.8 Using anthropometry in populations**

### **2.8.1 Uses related to decisions**

In populations, as in individuals, the major decisions for which anthropometric data are used relate to the types of intervention that are foreseen. Typical applications include decisions on whether or not intervention programmes are needed, to whom they should be delivered, and what their nature will be. These applications are similar to those involved in screening individuals; for populations, however, appropriate decisions are rarely as well established. Programme management, and timely warning and intervention systems to prevent famines and food crises (34, 35), for which population approaches have long been used, are probably exceptions to this general rule.

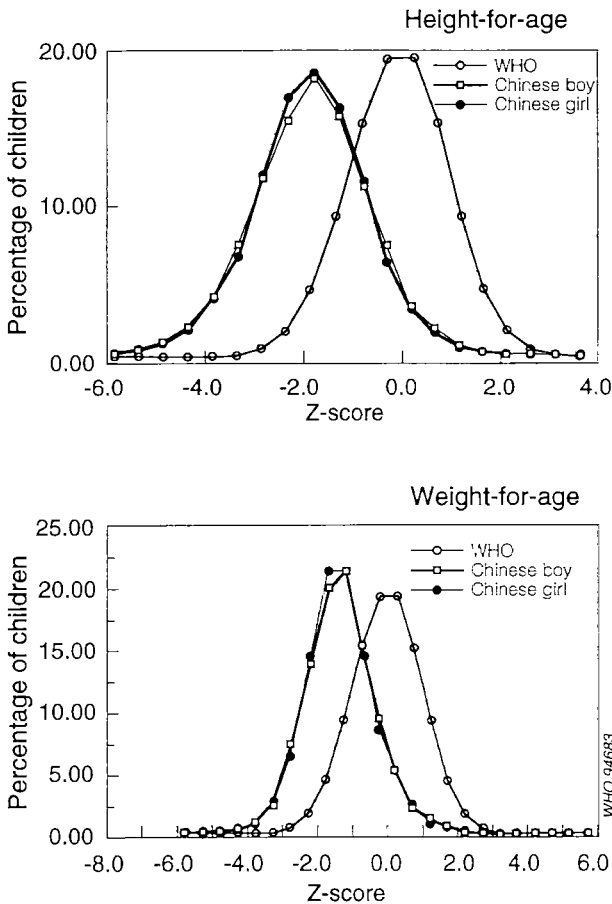
When the implementation of population interventions is planned, it is important to differentiate between relative risk and attributable risk (or, more specifically, the population attributable risk or etiological fraction). The risk of death in a child with a severe anthropometric deficit may be several times greater than that in a child with no deficit, while a child with mild deficits is at an intermediate level of risk. These comparisons refer to relative risks. In a population, however, the number of children with mild deficits will tend to be much greater than that of severely affected children. Thus, although severe deficits are associated with a larger relative risk, the mild deficits may account for the majority of deaths, which is the concept of attributable risk (24, 36). At the population level, its implication is that the overall impact of an

intervention will be limited if the intervention is delivered only to the most severely affected individuals.

A further important concept in the delivery of population interventions is that displacement of the whole anthropometric curve (Z-score distribution for the anthropometric indicator; see Fig. 7) often occurs in areas where nutritional problems are present. For example, data from many different countries show a very high consistency in the standard deviation of weight-for-height among young children expressed as Z-scores of the international reference. Even under conditions of extreme famine, where the mean Z-score is two or three units below the reference, the value of the standard deviation of Z-scores is very close to unity (38). This shows that the entire distribution is shifted, as seen in Fig. 7, so that

Figure 7

**Z-score distribution for height-for-age and weight-for-age of Chinese children compared with the NCHS/WHO international reference<sup>a</sup>**



<sup>a</sup> Source: reference 37.

all individuals, not only those below a given cut-off point, are affected (39). Interventions may consequently have to be directed at the whole population, rather than only at those individuals who fall below a given cut-off.

Appropriate use of anthropometry in populations must also take sampling strategies into consideration, including the choice of age ranges, time periods, geographical areas, and socioeconomic groups. Such technical issues as the relationship of sample size to statistical power, specific study designs, and confidence intervals are beyond the scope of this report and are treated elsewhere (6, 40, 41). The summary tables of appropriate uses of anthropometry provided in subsequent sections for different age and status groups give initial guidance on some variables to be considered for sampling purposes; however, specific expertise in sampling should be sought before surveys are launched, to ensure that the most important questions can be answered.

Considerations of sample size often result in the need to pool children of different ages, but this procedure is justified only if observed deviations from reference data have the same meaning relative to an intervention at different ages. For example, the cumulative effects of stunting may have ceased by the age of 3 years, so that prevalence values over a wide age range may be difficult to interpret. Assessment of stunting among older children, in whom there is a fixed deficit, will then yield more easily interpretable information regarding the need for long-term intervention. Because the intervention is to be directed towards young children, in whom there is active stunting, rather than towards those who are already stunted, this approach seems paradoxical. Nevertheless, older children provide the sentinel signal for the population to be targeted, even though they will themselves no longer benefit from the intervention. By concentrating monitoring on these “sentinel” children, information can be collected earlier and more cheaply, will be more understandable, and will have greater relevance to decisions regarding actions with longer-term impact.

Sampling considerations must include the appropriate timing of surveys; this is particularly important in such contexts as the alleviation of seasonal food crises. Timing is also a critical aspect of any decision-making process based on anthropometry; the ability to meet deadlines for the collection, compilation, analysis, and presentation of data may be as important as any other consideration. Expertise in designing surveys that are timely is essential if the collected data are to be transformed into information that is relevant and useful for effective public policy and action.

### **2.8.2 Targeting interventions**

A screening tool can be used to estimate prevalence by counting the number of individuals in a population who fall below a given cut-off point. Anthropometric indicators can also be used to characterize the

status of a population: the mean Z-score, for example, will provide a more accurate estimate of poor anthropometric status than observed prevalence (38), thus reducing the sample size needed for a nutritional survey. In anthropometry, differences in means provide greater statistical power than differences in prevalence in discriminating across target groups (38). Examples of this approach are discussed in more detail in section 5.

Sometimes comparison of the whole population distribution (as shown in Fig. 7) is indicated, rather than just the mean Z-scores or the prevalence below a given cut-off point. In a recent report on a refugee group (42), for example, the death rate among the most severely malnourished was so high that the lower end of the distribution was truncated, leaving the mean hardly affected. A full discussion of appropriate determination of prevalence is included in section 2.8.5.

In principle, targeting of populations, as of individuals, can be based not only on a one-time measurement as discussed above, but also on repeated measurements.

### 2.8.3 **Assessing response to an intervention**

Assessing the response to interventions requires at least two measurements. If the intervention is likely to affect the anthropometric characteristics of the individual, it is usually more efficient to measure the same individuals twice than different individuals on two occasions, because of the smaller sample size needed to identify a change. In other circumstances, repeated measuring of the same individual makes little sense, especially where prevention of a given condition is the objective of intervention. In such cases, different individuals of the same age are measured to assess reduction in prevalence. It is then essential to take into account any factors that may distort comparability over time, such as selective migration.

The problem of regression to the mean has already been discussed in the context of repeated measurements in the individual. It is less well recognized that the same phenomenon will occur in populations selected for their low initial values, even if the second measurement is not taken in the same individuals as the first.

The same Z-score deviations from the reference data do not necessarily have the same meanings at different ages. It is therefore impossible to interpret change properly unless the effect of age is taken into account.

When the response to specific interventions in a population is monitored, the time delay before the chosen indicator shows evidence of change must be taken into account. For instance, months or years are required to assess the effect on birth weight of improved nutrition during pregnancy, but decades for improvements in birth weight through prevention of childhood malnutrition to become apparent.

#### 2.8.4 **Ascertaining the determinants and consequences of malnutrition**

In general, relating anthropometric indicators of malnutrition to the determinants or consequences of the condition in populations requires careful distinction between non-causal and causal associations. The exception is in targeting, for which causal and non-causal relationships can be equally useful. For example, stunting of an older sibling is a good targeting indicator even though it is not a direct cause of malnutrition in the younger child.

Efforts to infer causality from a single survey must take account of non-causal associations arising from coincident changes across different birth cohorts. For instance, in survey data collected at one specific time, literacy and physical stature in adults may show an inverse relationship with chronological age. The reason for this is that increases in both stature and literacy are the consequences of secular improvements in socioeconomic development that have affected younger adults. This cohort effect, which is a characteristic problem in surveys of older people, is discussed in more detail in section 9. The need for correct modelling of relationships between indicators and determinants and consequences has already been addressed.

#### 2.8.5 **Nutritional surveillance**

Anthropometry provides some of the most important indicators used in nutritional surveillance. The following classification of nutritional surveillance (34) is based on the different types of survey mechanisms and other procedures necessary to collect, analyse, and transfer information for use in making decisions that affect nutrition.

##### *Surveillance for problem identification and for policy and programme planning*

Prevalence estimates often play a pivotal role in the assignment of government priorities to health problems. True prevalence can be estimated from measured prevalence by taking sensitivity and specificity into account (43), and this is often done for specific diseases for which the sensitivity and specificity of a given indicator are precisely calculated through comparison with a clinical or pathological “gold standard”. For anthropometric indicators of nutritional status, however, the “gold standard”, that is, appropriate nutrition, cannot be measured directly; this presents a difficulty in nutritional surveillance.

The logic of an alternative method of estimating the true prevalence of malnutrition is based on the assumption of a universally applicable specificity distribution. This assumption is closely approximated in young children (44), for whom this distribution corresponds to the growth potential of all populations of young children in which there are no stunting or wasting factors, currently represented by the NCHS/WHO

reference data.<sup>1</sup> The observed distribution of children in any population is made up of those who have not been stunted or wasted and thus correspond to the NCHS/WHO reference data (specificity distribution), plus stunted and wasted children (sensitivity distribution). The prevalence of stunting and wasting is the ratio of the children in the sensitivity distribution to all children in the population; the sensitivity distribution is obtained by subtracting the specificity population from the total population (45). To the degree that other universally relevant reference data for other healthy conditions (e.g. healthy thinness) can be defined, these reference data can be used as the specificity distribution for counting the unhealthy (e.g. the overweight) in conditions other than childhood malnutrition.

Two new methods have been proposed for estimation of prevalence using the reference data as the specificity distribution (46, 47). The more recent of the two (47) implicitly takes account of the effect of prevalence itself on the results, and is more accurate. This method is also the simpler, does not require a computer, and has good precision when the mean Z-score of the malnourished population is low. In other cases, however, a graphical method (45) using the reference standards as the specificity distribution might have better precision.

Computer methods are available that do not depend on an external standard to define specificity (48). Because they take into account the small genetic differences between populations, they should be intrinsically more precise provided that the sensitivity and specificity distributions are Gaussian.

All the above methods make the assumption that only a proportion of the children are malnourished. Other methods (38), described in section 5, may be used if *all* the children in the population can be assumed to be malnourished.

It is a mistake to compare relatively precise estimates of malnutrition prevalence derived from anthropometry in young children with those of other diseases estimated less precisely or with estimates of malnutrition based on other indicators. Dietary information derived by using cut-offs is particularly misleading in this application (49).

In nutritional surveillance for policy development and programme planning it is crucial to identify those of the most important causal influences that are amenable to interventions. Clear differentiation between analyses designed to *identify* interventions and those designed to *target* interventions is essential.

---

<sup>1</sup> NCHS/WHO growth reference data: reference data for height and weight of US children, originally collected by the National Center for Health Statistics and recommended by WHO for international use (see *Measuring change in nutritional status*, Geneva, World Health Organization, 1983).

### *Timely warning and intervention systems*

Child anthropometry has provided appropriate indicators for targeting food distribution to prevent outright famines. In order to obviate the need for the disruptive social and economic effects of food relief, timely warning and intervention systems should also be capable of averting food crises. Unfortunately, changes in the prevalence of wasting large enough to give reliable warning usually occur too late to permit effective preventive action against food crises (35); a longer lead time is essential if the need for emergency food distribution is to be avoided.

The timeliness of information can be improved by preferential sampling of “sentinels”, i.e. groups and individuals who signal the advent of food crisis earlier than the rest of the population. However, this early warning may lack specificity, in that it may occur seasonally even when there is no subsequent food crisis in the population. It therefore becomes necessary to follow this first screen with the collection of other more specific information (34).

Anthropometric indicators are useful in the late stages of the evolution of food crises (35). As famine progresses, however, selective mortality of the most wasted children may make the affected population as a whole appear less severely malnourished than it really is (42), and anthropometry must therefore be complemented by mortality information.

### *Surveillance for programme management*

Programme managers require information both for targeting an intervention and for evaluating its success: its efficacy in covering everyone it should, and its efficiency in covering *only* those that it should. These latter two aspects have very different implications for sampling. Efficiency, or yield, can be determined by assessing the positive predictive value on programme participants themselves; assessment of coverage depends on determining sensitivity and thus also requires information on non-participants, which presents a much more difficult task.

Programme managers also need to ensure that the response of participants is as expected. In defining expected responses in situations where participants are screened by anthropometry, it is important to take into account the positive predictive value and regression to the mean.

The impact of the intervention programme on participants is a matter of concern, too, to those who provide the financial resources, who often fail to realize that this cannot be assessed on the basis of data obtained from participants alone. Assessment requires appropriate control groups, and expertise is also needed to model the anthropometric impact properly. Understanding this model is essential both for sampling the individuals for measurement and for analysis of the data.

## 2.9 Characteristics of reference data

A *reference* is defined as a tool for grouping and analysing data and provides a common basis for comparing populations; no inferences should be made about the meaning of observed differences. A *standard*, on the other hand, embraces the notion of a norm or desirable target, and thus involves a value judgement. Concern has been expressed that, because reference data embody certain characteristics or patterns of normality, they have been widely and inappropriately used to make inferences about the health and/or nutrition of individuals and populations; that is, they have been treated as optimum targets, or standards, and any deviation from these “standards” has been assumed to have a fixed and particular meaning. Much of the justification for this is provided by extensive evidence that, in populations, the effect of ethnic differences on the growth of children is small compared with environmental effects. Thus, for example, there is no reason to believe that the 2–3 cm difference in median height between well nourished 18-year-olds in the Netherlands and France has any health implications, nor that improving the health and nutrition of French youth would be associated with any reduction in the height difference. By contrast, the Expert Committee recommended a body mass index cut-off of  $\geq 30$  as a provisional standard of grade 2 overweight (defined in section 7.2.1), applicable to *all* adults, because available data on risks of morbidity and mortality support this. For many other anthropometric characteristics, however, there are insufficient data to permit the specification of standards.

The Expert Committee recognized that release of references by WHO makes it almost impossible to prevent their use as standards for judging the nutritional status of individuals and populations. It is therefore recommended that care should always be taken to choose references that resemble, as far as possible, true standards, so that the same deviation from the reference data has the same biological meaning. For example, because the mean heights of young children from many affluent populations differ little across ethnic groups compared with the socioeconomic variability within a given ethnic group (44), it should be possible to construct a standard that represents the growth potential of all children. This may seem surprising in the light of the rather broad distribution of attained heights and weights, generally felt to be of genetic origin, within a well nourished population. However, across most populations there seems to be very little difference in mean growth in height or in its distribution around the mean that is attributable to genetics. A universal standard of height distribution among young children is therefore justified, but it must derive from a population that has fully met its growth potential. For this reason, the most important criterion in choosing the current set of WHO childhood height and weight reference data (4, 6) was that it should come from a well nourished population (50).



The choice of a sample for developing references or standards thus raises the question of what constitutes a healthy population. At least four definitions exist:

- (1) The population lives in a healthy environment. This is the type of population from which the current childhood NCHS/WHO reference data (4, 6, 7) have been drawn.
- (2) The population lives in a healthy environment and contains no overtly sick or very few clinically sick individuals. This is the type of population from which many national paediatric reference data have been drawn.
- (3) The population lives in a healthy environment and contains only individuals whose present good health will be demonstrated by longevity or at least by survival for some years after measurements are taken.
- (4) The population lives in a healthy environment and contains only individuals who live healthily according to present prescriptions, for example infants who are breast-fed according to WHO recommendations.

A further definition might cover some combination of the above, such as a population living in a healthy environment, excluding both those who die within some specified time after measurement (see item 3 above) and those who engage in unhealthy practices such as smoking (see item 4).

The first of the above definitions prevailed for children in the past because its advantages, principally total population representation, were felt to outweigh the advantages of the second, especially since the NCHS/WHO reference data were very similar to the best of previous reference data sets based on the second definition. Little work has been done in comparing populations that correspond to the first two definitions with others; section 5 of this report, however, contains a comparison of definitions (1) and (4) in the context of infants who are exclusively breast-fed from birth to 4-6 months of age in accordance with WHO recommendations (51).

A related question concerns the extent to which different standards should be used to approximate the ideal, i.e. whether there should be different standards of birth weight according to race, of growth data according to parental size, or of body mass index according to body frame. Possibly, this should depend upon the use to which the standards will be put. For instance, different criteria for assessing mean birth weight according to maternal smoking status might be useful: equal degrees of intrauterine growth retardation have different prognoses for children of smokers and non-smokers. However, in assessing the prevalence of intrauterine growth retardation, controlling for a mother's smoking would be wrong as it would mask an important problem. The

theoretical advantage of using different standards for specific purposes may thus be counterbalanced by equally strong theoretical disadvantages. For this reason, when reference data are to be used to make decisions about populations, it is better to use statistical methods to control for differences (such as those associated with different altitudes) within or across populations than to use different standards.

On the level of the individual, different standards have been proposed to take into account intrinsic differences in the expected optimal size associated with, for example, differences in altitude, parental heights, or feeding practices (whether a child is exclusively breast-fed or not). The utility of developing different standards for screening individuals depends essentially on the prevalence of the condition being screened. Unless its variability is low, prevalence has such a large effect on the positive predictive value that errors arising from the lack of separate standards are of little practical significance. In wealthy countries, however, where prevalence is low and therefore shows little variability, the use of separate standards might be justified. In such settings, the positive predictive values of anthropometric screens are so low that further screens generally become necessary. The trade-off between the cost of using multiple standards and the savings made by avoiding further screens remains to be investigated, but computerization of expected optimal growth on the basis of various characteristics of individuals will probably favour the use of individualized standards in wealthy populations. In poorer populations, where the use of different standards does not improve screening and poses considerable managerial problems, single standards should continue to be used.

If reference data are to be used as standards, the criteria for the reference population are of critical importance. The following criteria have been established as desirable, and are briefly reviewed (50):

- *“The sample should include at least 200 individuals in each age and sex group”*

This criterion relates particularly to the precision with which extreme percentiles or Z-scores are calculated. A sample of this size would provide the 5th percentile with a standard deviation of about  $\pm$  a 1.54 percentile, and is considered acceptable for individual-based applications (such as screening). In population-based applications, the sample size is also sufficiently large to allow differentiation between environmental and genetic effects on growth (52).

- *“The sample should be cross-sectional since the comparisons that will be made are of a cross-sectional nature”*

This is no longer considered essential, since longitudinal data can be presented cross-sectionally with minor adjustments. On the contrary, growth charts derived from cross-sectional data should not be used to monitor longitudinal data (31). Where several measurements are made

in the same individual, the slope of the line joining successive points on the growth chart is a direct measure of growth velocity. If the slope differs substantially from that of the neighbouring percentile curves, so that the data appear to cross percentiles, this is taken to be an indication of abnormal growth. However, since percentiles are derived from cross-sectional data and are relevant only to single measurements, their application to the interpretation of longitudinal data is inappropriate. This is particularly true during infancy and puberty. Correct interpretation of percentile crossing requires a different set of percentiles, derived from longitudinal data (31). Unfortunately, cross-sectional references continue to be widely misused for the interpretation of longitudinal data.

- *“Sampling procedures should be defined and reproducible”*
- *“Measurements should be carefully made and recorded by observers trained in anthropometric techniques, using equipment of well tested design and calibrated at frequent intervals”*

References should also include data on reliability and precision (as is true of the current NCHS/WHO childhood reference) (53). Both inter-observer variability and instrument error should be documented, and it is useful, though not essential, to have separate estimates of within- and between-observer components of reliability (54).

Where missing data have had to be “imputed” – that is, generated by means of a statistical algorithm based on a number of assumptions – they should be separately identified and the method by which they were derived should be clearly documented. Any “cleaning” procedures used to remove patently spurious data should also be described.

- *“The measurements made on the sample should include all the anthropometric variables that will be used in the evaluation of nutritional status”*

The various measurements taken from a single individual should be compared with reference data derived from a single population. This avoids the inconsistencies that may arise from using several different references for different measurements, such as weight and arm circumference.

- *“The data from which reference graphs and tables are prepared should be available for anyone wishing to use them, and the procedures used for smoothing curves and preparing tables should be adequately described and documented”*

There have been many recent developments in techniques for smoothing curves, which have implications for required sample sizes and accurate representation of the data.

## References

1. **Beaton G 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).
2. **Habicht J-P, Pelletier DL.** The importance of context in choosing nutritional indicators. *Journal of nutrition*, 1990, **120** (Suppl. 11):1519-1524.
3. **Wang ZM, Pierson RN Jr, Heymsfield SB.** The five-level model: a new approach to organizing body-composition research. *American journal of clinical nutrition*, 1992, **56**:19-28.
4. **WHO Working Group.** Use and interpretation of anthropometric indicators of nutritional status. *Bulletin of the World Health Organization*, 1986, **64**:929-941.
5. **Armitage P, Berry G.** *Statistical methods in medical research*, 2nd ed. Oxford, Blackwell, 1987.
6. *Measuring change in nutritional status: guidelines for assessing the nutritional impact of supplementary feeding programmes.* Geneva, World Health Organization, 1983.
7. **Dibley MJ et al.** Development of normalized curves for the international growth reference: historical and technical considerations. *American journal of clinical nutrition*, 1987, **46**:736-748.
8. **Dibley MJ et al.** Interpretation of Z-score anthropometric indicators derived from the international growth reference. *American journal of clinical nutrition*, 1987, **46**:749-762.
9. **Gorstein J et al.** Issues in the assessment of nutritional status using anthropometry. *Bulletin of the World Health Organization*, 1994, **72**:273-284.
10. **Esrey SA, Casella G, Habicht J-P.** The use of residuals for longitudinal data analysis: the example of child growth. *American journal of epidemiology*, 1990, **131**:365-372.
11. **Gorstein J.** Assessment of nutritional status: effects of different methods to determine age on the classification of undernutrition. *Bulletin of the World Health Organization*, 1989, **67**:143-150.
12. **Backett EM, Davies AM, Petros-Barvazian A.** *The risk approach to health care, with special reference to maternal and child health care, including family planning.* Geneva, World Health Organization, 1984 (Public Health Papers, No. 76).
13. **Brabin B.** An assessment of low birthweight risk in primiparae as an indicator of malaria control in pregnancy. *International journal of epidemiology*, 1991, **20**: 276-283.
14. **Monteiro CA, Torres AM.** Can secular trends in child growth be estimated from a single cross sectional survey? *British medical journal*, 1992, **305**:797-799.
15. **Shekar M, Habicht J-P, Latham MC.** Is positive deviance in growth the converse of negative deviance? *Food and nutrition bulletin*, 1991, **13**:7-11.
16. **Brownie C, Habicht J-P, Cogill B.** Comparing indicators of health or nutritional status. *American journal of epidemiology*, 1986, **124**:1031-1044.

17. Habicht J-P, Meyers DL, Brownie C. Indicators for identifying and counting the improperly nourished. *American journal of clinical nutrition*, 1982, 35(5 Suppl.):1241-1254.
18. Galen RS, Gambino SR. *Beyond normality: the predictive value and efficiency of medical diagnoses*. New York, Wiley, 1975.
19. Leon ME et al. Identifying the malnourished within Peru: regional variation in the performance of a nutrition indicator. *International journal of epidemiology*, 1990, 19:214-216.
20. Briend A et al. Measuring change in nutritional status: a comparison of different anthropometric indices and the sample sizes required. *European journal of clinical nutrition*, 1989, 43:769-778.
21. Yarbrough C et al. Response of indicators of nutritional status to nutritional interventions in populations and individuals. In: Bosch S, Arias J, eds. *Evaluation of child health services: the interface between research and medical practice*. Washington, DC, U.S. Government Printing Office, 1978 (DHEW Publication No. (NIH) 78-1066):195-207.
22. Stoltzfus RJ et al. Evaluation of indicators for use in vitamin A intervention trials targeted to women. *International journal of epidemiology*, 1994, 22:1111-1118.
23. *Principles and practice of screening for disease*. Geneva, World Health Organization, 1968 (Public Health Papers, No. 34).
24. Last JM. *A dictionary of epidemiology*, 2nd ed. New York, Oxford University Press, 1988.
25. Victora CG et al. Evidence for protection by breast-feeding against infant deaths from infectious diseases in Brazil. *Lancet*, 1987, ii:319-322.
26. Habicht J-P, DaVanzo J, Butz WP. Mother's milk and sewage: their interactive effects on infant mortality. *Pediatrics*, 1988, 81:456-461.
27. Vecchio TJ. Predictive value of a single diagnostic test in unselected populations. *New England journal of medicine*, 1966, 274:1171-1173.
28. Victora CG et al. Is it possible to predict which diarrhoea episodes will lead to life-threatening dehydration? *International journal of epidemiology*, 1990, 19:736-742.
29. Fletcher RH, Fletcher SW, Wagner EH. *Clinical epidemiology: the essentials*. Baltimore, Williams & Wilkins, 1982:41-58.
30. Victora CG et al. Birthweight and infant mortality: a longitudinal study of 5914 Brazilian children. *International journal of epidemiology*, 1987, 16:239-245.
31. Cole TJ. Growth charts for both cross-sectional and longitudinal data. *Statistics in medicine*, 1994, 13:2477-2492.
32. Davis CE. The effect of regression to the mean in epidemiologic and clinical studies. *American journal of epidemiology*, 1976, 104:493-498.
33. Rivera JA, Habicht J-P, Robson DS. Effect of supplementary feeding on recovery from mild to moderate wasting in preschool children. *American journal of clinical nutrition*, 1991, 54:62-68.
34. Mason JB et al. *Nutritional surveillance*. Geneva, World Health Organization, 1984.

35. Brooks RM et al. A timely warning and intervention system for preventing food crises in Indonesia: applying guidelines for nutrition surveillance. *Food and nutrition*, 1985, 11:37-43.
36. Pelletier DL, Frongillo EA Jr, Habicht J-P. Epidemiologic evidence for a potentiating effect of malnutrition on child mortality. *American journal of public health*, 1993, 83:1130-1133.
37. Chang Ying et al. Nutritional status of preschool children in poor rural areas of China. *Bulletin of the World Health Organization*, 1994, 72:105-112.
38. Yip R. Expanded usage of anthropometry Z-scores for assessing population nutritional status and data quality. In: *Abstracts book no. 1, 15th International Congress of Nutrition (Adelaide)*. Adelaide, International Union of Nutritional Sciences, 1993:279.
39. Rose G. Sick individuals and sick populations. *International journal of epidemiology*, 1985, 14:32-38.
40. Lwanga SK, Lemeshow S. *Sample size determination in health studies: a practical manual*. Geneva, World Health Organization, 1991.
41. *Conducting small-scale nutrition surveys: a field manual*. Rome, Food and Agriculture Organization of the United Nations, 1990 (Nutrition in Agriculture, No. 5).
42. Yip R, Sharp TW. Acute malnutrition and high childhood mortality related to diarrhea. Lessons from the 1991 Kurdish refugee crisis. *Journal of the American Medical Association*, 1993, 270:587-590.
43. Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. *American journal of epidemiology*, 1978, 107:71-76.
44. Habicht J-P et al. Height and weight standards for preschool children. How relevant are ethnic differences in growth potential? *Lancet*, 1974, i:611-614.
45. Meyers LD et al. Prevalences of anemia and iron deficiency anemia in black and white women in the United States estimated by two methods. *American journal of public health*, 1983, 73:1042-1049.
46. Mora JO. A new method for estimating a standardized prevalence of child malnutrition from anthropometric indicators. *Bulletin of the World Health Organization*, 1989, 67:133-142.
47. Monteiro CA. Counting the stunted children in a population: a criticism of old and new approaches and a conciliatory proposal. *Bulletin of the World Health Organization*, 1991, 69:761-766.
48. Brownie C, Habicht J-P, Robson DS. An estimation procedure for the contaminated normal distributions arising in clinical chemistry. *Journal of the American Statistical Association*, 1983, 78:228-237.
49. *Nutrient adequacy: assessment using food consumption surveys. Report of the Subcommittee on Criteria for Dietary Evaluation of Food Consumption Surveys*, Food and Nutrition Board. Washington, DC. National Academy Press, 1986.
50. Waterlow JC et al. The presentation and use of height and weight data for comparing nutritional status of groups of children under the age of 10 years. *Bulletin of the World Health Organization*, 1977, 55:489-498.

51. The World Health Organization's infant-feeding recommendation. *Weekly epidemiological record*, 1995, **70**:119-120.
52. Martorell R, Habicht J-P. Growth in early childhood in developing countries. In: Falkner F, Tanner JM, eds. *Human growth: a comprehensive treatise. Vol. 3: Methodology: ecological, genetic and nutritional effects on growth*. New York, Plenum Press, 1986:241-262.
53. Marks GC, Habicht J-P, Mueller WH. Reliability, dependability, and precision of anthropometric measurements. The Second National Health and Nutrition Examination Survey 1979-1980. *American journal of epidemiology*, 1989, **130**:578-587.
54. Villar J et al. Perinatal data reliability in a large teaching obstetric unit. *British journal of obstetrics and gynaecology*, 1988, **95**:841-848.