66

prognostic difference between any of the groups—particularly no difference in age or time from primary diagnosis to first metastasis.

DISCUSSION

One might expect chemotherapy to lengthen overall survival in metastatic breast cancer because survival is longer from the time of start of treatment in patients who respond than in those who do not. However, despite high response rates we have seen no improvement in overall survival for patients with primary breast cancer, nor a prolongation of survival from first metastases. In fact, since the introduction of multiple-drug chemotherapy survival from first metastasis seems to have shortened.

It is possible that this reduction in survival reflects a change in the behaviour of the tumour, although this seems unlikely with the observed constant time from primary diagnosis to first metastasis seen with these two groups of patients. We appreciate that this is not a randomised controlled clinical trial, and bias may have occurred, although we have been unable to identify any differences in prognostic factors between the groups of patients.

The paradox of improved survival for responders to chemotherapy with apparent reduction in survival of the whole group since introduction of chemotherapy may arise in several ways. First, the good which is done to patients who respond to chemotherapy may be more than outweighed by the harm done to non-responding patients by non-effective, potentially toxic chemotherapy. Second, the patients who respond to chemotherapy may be those with less aggressive disease and therefore better prognosis. Survival of patients given chemotherapy is usually related to the time from start of treatment, which may vary from the time of first metastasis to the preterminal stage and therefore is meaningless in relation to overall survival. Finally, increasing use of chemotherapy may have reduced or delayed potentially successful endocrine therapy. We are now examining these factors in a larger series of patients.

We do not doubt that cytotoxic chemotherapy prolongs survival of some patients with metastatic breast cancer—particularly those with life-threatening rapidly developing metastases in lung, liver, or bone-marrow. We also have no doubt that this treatment may be of palliative benefit to many patients without necessarily prolonging survival. We suggest that these results indicate the need for more accurate identification of patients who will not benefit from potentially harmful cytotoxic treatment. Overall survival and quality of life may improve if these patients are not treated. We also suggest that existing analytical methods for assessment of response, depending on (1) measurement of objective regression and (2) differences in survival from start of treatment, are misleading and have encouraged a falsely optimistic view of cytotoxic chemotherapy for metastatic breast cancer. Assessment of palliation and differences in survival from first recurrence may be more informative.

The fact that regressions of breast cancer had no influence on overall survival must reflect the inadequacy of present-day chemotherapy. Clearly in the future with more effective treatments, such as we now have for Hodgkin's disease and some other tumours, objective regression may assume a closer relation to survival.

We thank Richard Skeete of the Cancer Registry and Clair Chilvers of the Department of Epidemiology at the Institute of Cancer Research for their help and advice. We particularly thank Caroline Gordon, our research assistant, for her help with this project, and Dr T. J. McElwain for his helpful criticism of the script.

This study was supported, in part, by the Breast Cancer Research Trust.

Requests for reprints should be addressed to T. J. P., Royal Marsden Hospital, Sutton, Surrey SM2 5PX.

REFERENCES

1. Bambilla C, DeLena M, Rossi A, Valagussa P, Bonadonna G. Response and survival in advanced breast cancer after two non-cross-resistant combinations. *Br Med J* 1976; i: 801–04.
2. Carter SK. Single and combination nonhormonal chemotherapy in breast cancer. *Cancer* 1972; **30**: 1543–55.
3. Russell JA, Baker JW, Dady PJ, et al. Combination chemotherapy of metastatic breast cancer with vincristine, adriamycin and prednisolone. *Cancer* 1978; **41**: 396–99.
4. Hayward JL, Rubens RD, Carbone PP, Heuson J-C, Kumaoka S, Segaloff A. Assessment of response to therapy in advanced breast cancer. *Br J Cancer* 1977; **35**: 292–98.
5. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomised clinical trials requiring prolonged observation of each patient: II analysis and examples. *Br J Cancer* 1977; **35**: 1–39.
6. McElwain TJ. The management of Hodgkin's disease. In: Turner P, ed. Advanced medicine: topics in therapeutics. London: Pitman Medical, 1976: 65–77.

Child Care

ECOLOGICAL CONSIDERATIONS IN THE CREATION AND THE USE OF CHILD GROWTH STANDARDS

H. GOLDSTEIN J. M. TANNER
Institutes of Education and Child Health, University of London, London

Summary There is no proper substitute for a country, especially a developing country, having its own child growth standards or norms for clinical use, based on a representative sample of the population. Separate standards may be derived for subgroups of the population, but the application to the whole population of standards based on an economically privileged group is inappropriate, as is the use of an international standard. The screening or clinical use of growth standards should be sharply distinguished from the use of growth measurements to compare disadvantaged with privileged groups or populations. In particular, the use of growth standards to screen individual children should not divert attention from the need to change existing differences between disadvantaged and privileged groups.

INTRODUCTION

The physical growth and development of children is a sensitive indicator of the health of a population. But how measurements of growth should best be used for this purpose is still not agreed. Are so-called standards essential, and if so from what sort of sample should they be derived and how should the results be interpreted?[1]

There are two distinct uses of growth measurements. First, they may be used to assess differences in health and nutrition between groups in a population and to

monitor changes in such groups or in the whole population over time. This involves the selection and measurement of representative samples of the groups or populations concerned and, typically, comparison of mean values between populations and times. Secondly, they may be used as a screening device to assess whether an individual child is abnormal, in the sense that his measurement lies beyond an extreme percentile of a population distribution (e.g., whether a child's height falls below the third percentile of height for all children of the same age). We are here chiefly concerned with the latter use, that is, with growth standards (or norms), although the two uses are closely connected.

With large samples it is easy to define accurate standards for different groups, such as those in urban and rural areas, or those with different ethnic backgrounds. We can assess a child according to the standards for his own subgroup rather than those for the whole population. Thus, a child from a deprived region could be assessed according to standards for that region. It is open to question if this is desirable. A child in such a region should perhaps be judged by the standards for an economically privileged area since these indicate the growth potential of the child who, like other children in his region, has failed to realise that potential because of various economic and other environmental circumstances. This would apply particularly in countries where there are large growth differences between rich and poor. We shall discuss the basis for this proposal and make some suggestions for the efficient use of growth standards with particular reference to developing countries.

BACKGROUND

In 1971 a committee of the International Union of Nutritional Sciences[2] made specific recommendations for the establishment of growth standards. They said:

"The Commission strongly recommends that studies be carried out in as large a variety of countries as possible. Each country's own standards must be derived from carefully selected samples representing children growing in an optimal environment for that country".

"In selecting the specific population the first group should be from the 'modern elite' groups in each study area".

"It is felt that justification for the creation of growth standards is as follows. Anthropometric measures are the most important means of assessing nutrition and health in communities, especially in children. Furthermore, appropriately developed standards can serve as a reference against which to measure changes in health and nutrition of a given country and also as standards for evaluating the results of intervention programmes".

Other authors have supported these recommendations[3] although they have not gone unchallenged.[4] The quotations illustrate clearly some of the confusion prevalent in this field. Firstly, there is a failure to distinguish the different uses of these standards. For example, "élite" standards are obviously inappropriate for assessing whole population changes in health and nutrition or for evaluating the results of intervention programmes. Secondly, the definition of an élite group is ambiguous. The quotation mentions optimal environments, but does not define optimal. This problem has been described thus:[5]

"There is, all the same, one argument against using "best-off" standards. The best-off in most countries grow up earlier and end up taller. Is early maturation and large size advantageous? . . . Growth is indeed a fine yardstick of the health of individuals and populations, perhaps the best there is. But it remains so only for as long as we view our standards as a sensitive balance to be adjusted if conditions change, and not an immutable ceiling to which we should all eat our way."

Unless optimum can be defined and be shown to be operationally useful and valid the argument for the general clinical use of privileged-group standards largely disappears. In the next section we shall explain why this is so.

STANDARDS BASED ON ECONOMICALLY PRIVILEGED GROUPS

The argument in favour of using privileged-group standards runs somewhat as follows. In a given country it may be reasonable to assume that all individuals have a common gene pool. There are some individuals who constitute an economically privileged group within that society and whose standards of nutrition, medical care, and so on, are better than those of the rest of the population. The environment of these individuals is said to be "optimal". These individuals, therefore, should constitute the standardising group. In this argument, however, the term optimal is used in a sense we think mistaken, namely that it is associated with the group living in the most "sophisticated" or technologically advanced environment. Although, in the recent past, the best nutrition has often been equated with the most food that can be obtained, it is now accepted that too much food may be as harmful as too little. A better definition of optimal would be the level of nutrition and medical care which is associated with the greatest amount of health. Since, unfortunately, there is not yet a satisfactory definition of positive health, we normally have to use a criterion based on the lowest mortality and morbidity rates.

Those who advocate the use of privileged standards often do so because they believe that all individuals in a population have a right to attain the status of those who are in the economically most privileged group. While this argument has the appearance of being socially progressive, it is in reality a superficial appearance. We think it only valid if it is applied to *groups* rather than individuals (and if "most healthy" is substituted for "most privileged").

The point is that, *within* a given environment, the minimum morbidity is not necessarily associated with the maximum rate of growth. Many factors, including physical activity, climate and culture, as well as general economic level, interact with growth processes to determine what is optimal, in terms of minimum morbidity and mortality rates. Thus, in a poor environment, a child who is small may have an advantage over faster-growing children in terms of mobidity and mortality *in that poor environment*. He remains disadvantaged by comparison with a child from the privileged environment because of the *overall* differences in morbidity and mortality rates between the two environments. Thus though we may describe an individual as performing as well as he can *for his environment* we must at the same time explicitly recognise the poverty of that environment itself. In these circumstances, therefore, it is the environment itself that needs altering, not just the circumstances of the individual child. As the environment of the disadavantaged group is improved, and mobidity

and mortality rates approach those of the privileged group, so the definitions of optimal growth in the two groups will tend to coincide.

Advocates of the use of privileged-group standards for clinical assessment of total populations argue that *each individual's environment* should be suitably altered if his or her development is judged to be suboptimal in relation to these standards. They point out that since environments are not completely homogeneous, and even in poor environments some individuals may not be very badly off, it may be more efficient to treat only those particular individuals identified as suboptimal in this way. Treatment might involve nutritional supplementation or more frequent medical attention. There are, however, practical objections to this argument. It is difficult to alter one individual's immediate environment without affecting others, and it is often socially unacceptable to do so. Moreover, such measures usually do little to remedy the underlying reasons for the poor environment itself, since they are not primarily aimed at raising the general economic level of its inhabitants. Although such a programme may be initially successful, if the extra nutritional or health interventions are subsequently withdrawn or reduced, the individual may be no better off than before. Indeed he may be worse off as a result of having to revert to his former level of subsistence. Thus, such interventions are useful only if they imply a continuing commitment to the alteration of the environment.

If we accept that we should design programmes to alter the general environmental conditions, growth measurements will still be necessary to establish the existence and size of growth differences between groups or populations. Such studies do not necessarily require percentile standards, but they must be based on proper sampling and measurement procedures to give reliable estimates of average differences. Without such basic knowledge about the total population, the widespread use of privileged standards may merely divert attention from the more urgent task of measuring and then dealing with overall health inequalities.

Nevertheless, for any given environment it is no doubt useful to have percentile growth standards to screen for abnormalities that respond to medical, social, or nutritional treatment. For example, there will always be children considered too small who may need clinical attention to decide whether any specific treatment is necessary. In a poor environment such measures will be additional to those done to improve the environment overall. The percentile standards used should be those of the population of the environment which actually exists, rather than those of a privileged group, and should be updated as often as necessary. The purpose of percentile growth standards is to screen *individual* children in relation to other children; the purpose of comparing average values of *groups* is to identify those groups which may require a reallocation of social, medical, or other resources to alter their general environmental conditions.

Finally, we have briefly to query the assumption that the privileged group has the same genetic background as the remainder of the population. There are relatively few developing countries that are genetically homogeneous in the sense that economic divisions do not also partly reflect ethnic divisions. Indeed, this is true also in many industrial nations. Until the distribution of particular genes in groups of the population is better understood, it is difficult indeed to allow for this factor.

IMPLICATIONS

We have argued that the case for applying privileged-group standards to a whole population is dubious. How many separate standards should then be used? Clearly it is inconvenient to have several hundred each applying to a different subgroup. However, men and women are nearly always considered separately in the population, because the sexes develop differently. Standards for children's growth based on the height of their parents are increasingly used[6] and this seems a useful approach towards coping with genetic and ethnic differences. Such standards enable a greater precision to be applied in determining whether a child is small in relation to others in the population. The same considerations about improving the whole environment still apply, of course, since part of the resemblance between parents and children may reflect similar environments. A similar situation also exists with respect to birthweight.[7]

In defining the subgroups of the population for which separate standards might be created we have first to consider the ease with which such children can be classified and secondly, to select groups of children that have large differences in growth, because this will increase the accuracy with which a child can be ranked. Sex, family size, where a child lives, and his ethnic characteristics are usually easy to classify. Parents' height is a more sensitive assessment than sex at most ages, but this must be to some extent counterbalanced by the ease with which sex can be determined.

In some developing countries it may be useful to have separate standards for different regions where there are large growth differences, which are updated frequently in circumstances where changing economic and social conditions are reducing these differences. The value of different standards for clinical and administrative use could thus be assessed. Such a procedure is being tried in Cuba, where as well as a national set of standards for use in the clinical assessment of all children, there will also be separate standards for experimental use in different areas so that their usefulness can be properly monitored.[8]

CONCLUSIONS

Percentile growth standards for screening children should be derived from the population or subpopulation to which the children belong. It may be inappropriate and even harmful to use standards derived from an economically privileged group. It is also inappropriate in one country to use standards derived from another. The concept of a single international growth standard is invalid, although there is some use for international reference values.* Imported standards, usually derived from populations of industrialised countries, are sometimes

*Such an approach has been developed by the W.H.O. in its growth chart for international use.[9] In this a series of reference values, derived from one set of measurements, permits a country or health service area to select a value range that most closely approximates to values of apparently healthy children in the population in question. Although this is evidently not an ideal approach it does offer a simple and practical instrument for health screening by primary health care workers.

better than no standards at all. But if standards do not exist in a particular developing country, rather than adopting an international standard it would be better to seek a set of standards that have been derived for a country with similar economic, social, and ethnic circumstances. This may not be easy, but it would be better to attempt it than to accept uncritically an inappropriate standard.

The comparison of whole populations between countries, and subpopulations within countries, is different, and is more rapidly achieved by comparing average growth measurements, rather than by preparing percentile standards.

We thank Prof. N. R. Butler, Prof. M. J. R. Healy, and Prof. J. Jordan and his colleagues at the Cuban Instituto de la Infancia for their very helpful comments.

Requests for reprints should be addressed to J. M. T., Institute of Child Health, 30 Guilford Street, London WC1N 1EH.

REFERENCES

1. Tanner JM. Human Growth Standards: construction and use. In: Gedda L, Parisi P, eds. Auxology: Human Growth in Health and Disorder. London: Academic Press, 1978.
2. International Union of Nutritional Sciences. The creation of growth standards: a committee report of a meeting in Tunis. *Am J Clin Nutr* 1971; 25: 218–20.
3. Habicht JP, Martorell R, Yarbrough C, Malina RM, Klein RE. Height and weight standards for pre-school children. *Lancet* 1974; i: 611–15.
4. Goldstein H. Height and weight standards. *Lancet* 1974; i: 1051–52.
5. Tanner JM. Population differences in body size, shape and growth rate: a 1976 view. *Arch Dis Child* 1976; 51: 1–2.
6. Tanner JM, Goldstein H, Whitehouse RH. Standards for children's height at ages 2–9 years allowing for height of parents. *Arch Dis Child* 1970; 45: 755–62.
7. Tanner JM, Thompson AM. Standards for birthweight at gestation periods from 32 to 42 weeks allowing for maternal height and weight. *Arch Dis Child* 1970; 45: 566–69.
8. Jordan J, Ruben M, Hernandez J, Rebelagua A, Tanner JM, Goldstein H. The 1972 Cuban national child growth study as an example of population health monitoring: design and methods. *Ann Hum Biol* 1975; 2: 153–71.
9. W.H.O. 1980 A growth chart for international use in maternal and child health care: guidelines for primary health care personnel. Geneva: W.H.O., 1978.

Occasional Survey

THE POSSIBILITY OF PREVENTING AMBLYOPIA

R. M. INGRAM

General Hospital, Kettering, Northamptonshire NN16 8UZ

INTRODUCTION

AMBLYOPIA is a reduced visual acuity, unilateral or bilateral, in eyes which are organically healthy. It was once thought to be caused by squint, but is frequently found in the absence of noticeable squint. Squint and/or amblyopia arises in about 7% of children, so these disorders are much seen in ophthalmic clinics.[1,2] Of the various treatments,[3] occlusion is the best known, but doubts persist about its efficacy.[4–11] Reported results are not satisfactory[12–19] and many children relapse.[20,21] Tour observed that, "after two centuries of using occlusion, we still do not know which eye to occlude, with what or for how long".[22] Attempts have also been made to stimulate the amblyopic eye with pleoptics[21,23] or rotating gratings,[24] but neither completely restored visual acuity. Non-strabismic amblyopia does not consistently and permanently improve with occlusion.[19] The numerical importance of this type of amblyopia has only lately been recognised.[25,26] Formerly, ophthalmologists dealt mainly with amblyopia associated with squint, and in such cases the most severe amblyopia undoubtedly does improve although probably only to a level of vision that was present at the time the squint appeared (as Chavasse apparently suspected[27]).

VISUAL-DEPRIVATION STUDIES IN ANIMALS

Some understanding of the basic factors in squint and amblyopia has come from investigation of the developing visual system of kittens and infant monkeys. These studies, started by Hubel and Wiesel in 1963,[28] have shown that visual deprivation—i.e., blurring of vision by various methods—during a "sensitive period", the first three to four months of life, has a permanent detrimental effect on the neuronal connections between retina and occipital cortex. The result is decreased acuity of the visually deprived eye and loss of binocular function. Switching occlusion to the good eye allows recovery of vision in the deprived eye, but only if the switch takes place within the sensitive period,[29] and even then the recovery is at the expense of binocular function.[30] Visual deprivation (e.g., occlusion) after the sensitive period neither creates nor corrects amblyopia.

CLINICAL IMPLICATIONS OF THE SENSITIVE PERIOD

It would be strange if there was not a corresponding period in man within which visual deprivation had a similar effect on vision. Some evidence for this comes from Japanese children who had operations for entropion. Postoperatively the eye was covered for one week; and all those treated before eighteen months of age ended up with amblyopia and defective binocular vision.[31] In the U.K., Ikeda and Wright[32] warned that, if one eye of a child was occluded during the sensitive period, the occluded eye could become amblyopic. This worried orthoptists, who forgot that such a catastrophe seldom happened in practice. If occlusion does not regularly create amblyopia, why should we expect the neuronal connections of the amblyopic eye to be amenable to alteration?

The observation that occlusion did not permanently improve the vision of the non-squinting amblyopes, or of squinting amblyopes whose squint was of short duration, led to the suggestion that visual deprivation in infancy had caused abnormal neuronal connections between retina and cortex,[19] and that these connections were unalterable because the children had presented after the sensitive period.[33] They[19] called this *primary* amblyopia, and suggested that the more severe amblyopia associated with longstanding squint, which does partly respond to occlusion, might have a different origin and be, for some reason, *secondary* to the "binocular interaction" of prolonged squint. This agrees with Chavasse's dual classification—(1) amblyopia of arrested development; and (2) amblyopia caused by "disuse"—and is similar to Abraham's classification.[34]

Earlier this century, ophthalmologists advocated occlusion to "prevent" amblyopia; but it is doubtful if they really prevented anything more than, in some children, the onset of secondary amblyopia.