

What is normal?

J. H. KOESLAG

Abstract The distinction between normality and abnormality forms the basis of medical practice. However, these words appear to have no more precise meanings in medicine than they do in conversational English. At least five contradictory definitions are described in the literature, and are in simultaneous use in everyday practice. The apparently arbitrary manner in which these definitions are chosen to evaluate different phenomena effectively means that medicine operates without a definition: certain phenomena are normal (by decree, as it were), and others are not. Actions based on such an arbitrary system are, of necessity, haphazard. The adoption of a precise, rigorously acultural definition of normality would unquestionably admit medicine to full membership of the family of objective sciences.

S Afr Med J 1993; 83: 47-50.

The categorisation of biological phenomena as normal or abnormal is the foundation of medical practice. Without it there could be no health services. It is therefore of fundamental importance to know on what this dichotomy is based. Yet none of the standard physiology and medical textbooks provide a clear and unambiguous definition of normality. *Dorland's Medical Dictionary*,¹ before describing some of the jargon associated with the word in chemistry and bacteriology, defines normal as 'agreeing with the regular and established type'. It does not, however, indicate what the 'regular or established type' is, or how it is determined. Scadding²⁻⁵ discusses the converse concept and defines a disease as: '... the sum of the abnormal phenomena displayed by a group of living organisms in association with a specified common characteristic or set of characteristics by which they differ from the norm of their species in such a way as to place them at a biological disadvantage'. This merely replaces the word 'disease' with the words 'abnormal', 'differing from the norm', and 'disadvantageous', each of which can have several meanings. Most physiology and medical textbooks provide lists of 'normal values' without, however, stating the criteria by which they are judged to be normal.

Clearly treatment should be directed primarily at the abnormal. If therapy is aimed at altering normality (into abnormality) then it must be applied with the greatest circumspection to achieve a clearly defined, and carefully considered, objective. An example is fertility control, where an unquestionably normal physiological phenomenon is deliberately rendered unoperational. Under most ordinary circumstances, however, the iatrogenic conversion of normality into abnormality (as a primary objective) is anathema. In other words, it is as important for the medical profession to remain on the correct side of the normal/abnormal boundary as it is for motorists to keep to the correct side of the road. In both cases the boundary may on occasions be crossed, but only with full awareness of the implications and risks.

Adventures on the wrong side of the boundary as a result of ignorance are clearly unacceptable and dangerous.

Definitions of 'normal'

It appears that the term, 'normality', has the same meaning in physiology and medicine as it does in conversational English. In all three it denotes at least five independent, mutually contradictory states: (i) the usual; (ii) not ill; (iii) operating as intended; (iv) conforming to a cultural norm; and (v) the best.

Special uses of the word 'normal' (often with a capital N) to denote perpendicularity in geometry, equivalence in chemistry, isotonicity with red blood cells in fluid therapy, and co-operation in politics, seldom cause confusion in medicine and will therefore be ignored in this discussion. The jargon use of the word 'normal' to denote a Gaussian distribution in mathematical statistics unfortunately does cause confusion,⁶ though unnecessarily so. Some biological data happen to fit a Gaussian distribution allowing the use of Gaussian statistics for their description and analysis. Most biological data, however, fit other distributions. Different mathematical tests must, therefore, be employed to analyse these data (e.g. a non-parametric test). However, these distributions and tests make no statement about the biological normality of the information.

The usual

Normal is synonymous with 'usual' in all the physical sciences and in much non-human biology. It is thus normal for bears to sleep continuously during winter and female spiders to eat the males that mate with them.

In human biology 'normal' is sometimes also defined as the usual. Thus the normal height of young men is the range of heights that includes 95% of men. It is also normal for newborn babies to sleep for more than 20 hours per day. It is similarly normal for non-pregnant young women to experience intermittent uterine bleeding.

In broad terms, the unusualness of a symptom is possibly the single most important determinant of whether a patient seeks medical advice or not. Thus, it is normal to feel short of breath after climbing five flights of stairs. But shortness of breath after brushing one's teeth is unusual, and therefore a good reason to consult a physician.

This usage of the term 'normal' is based on strict observation. It is rigorous and acultural. It dispassionately describes the species exactly as it is. The 95% probability limits constitute the most widely accepted normal range. This coincides with the internationally accepted limits of statistical significance (i.e. extra-normality is denoted by $P < 0,05$). As such, it constitutes an indispensable tool in the diagnostic industry, a point that will be elaborated in the discussion.

Not ill

This would appear to be the most useful of the medical definitions of normality.²⁻⁶ However, it is based on a circular argument or tautology. 'Illness' cannot easily be defined independently of the term 'normal': it is, by common acknowledgement, the state of being abnor-

mal.⁵ In other words, 'normal', by this definition, means 'not abnormal'.

In theory it would be possible to compile a catalogue of 'illnesses' the exclusion of which defines a person as being normal.^{5,6} But on what grounds, other than by decree or tradition, would an item qualify for inclusion or exclusion in that catalogue? Clearly a phenomenological definition of illness would have to be devised. Illness might thus be defined as anything that causes incapacitation, pain, swelling, bleeding, or infection.⁵ But is sleep then abnormal? Would child-birth constitute ill-health? Is an erection a surgical emergency? Would menstruation be a sickness? Are acne and colonic fermentation diseases?

This definition of normality is cumbersome and subjective. Its application to human behaviour is particularly likely to become bogged down in a quagmire of cultural prejudice. Scadding's⁵ contention that this subjectivity disappears when the aetiology becomes known, is the result of a confusion between the 'definition of disease' and the 'description of individual diseases'. A perfect understanding of the precise hormonal mechanisms of menstruation, for instance, does not contribute to the evaluation of this process as normal or not. The McMaster Group⁶ appears to be unaware of the tautologous nature of this definition, and provides no avenue of escape.

An objective, alternative definition of 'ill' is presented in the discussion.

Operating as intended

If the purpose of sex is procreation, then homosexuality and masturbation are, by this definition, deviant behaviours. The bottle-feeding of babies, for similar reasons would also be considered abnormal.

This definition of normality is used particularly forcefully in the fields of nutrition, exercise physiology, and sexology. It assumes that the designer's intentions are known, and, more importantly, contends that these intentions may not be defied. This is indeed the great weakness of this definition. It is the business of science to unravel the mechanistic reasons for natural events. The teleological reasons are essentially beyond the scope of science. Scientists and philosophers might speculate about them, but they can never be known.^{7,8} For example, what is the purpose of the hair on our heads, and is it normal to cut or style it?

The dangers inherent in this definition are obvious. The boundary between normal and abnormal is based on speculation which is inescapably ethnocentric, and often based on simplistic biological notions.^{7,9} It presupposes, furthermore, a static creationist origin of nature in which the creator, whether a deity or the process of natural selection, after moulding perfection during 6 symbolic days of intensive creativity, has been completely at rest for the past 10 000 or possibly a million years. If creation/evolution is an ongoing process then nothing has an immutable purpose or function, and normality can only be described in terms of current usage.⁹

If it is abnormal to use the genitalia for anything but procreation, then it is presumably also abnormal to use the tongue to lick postage stamps, the eyelids to wink at a friend, the vocal cords for speech (since they were originally presumably intended only for cries of alarm and grunts of pleasure), or the legs for purposes other than walking.

Conforming to a cultural norm

The normality of human behaviour is to a large extent judged by this criterion.⁶ An aversion to jogging, a love of rich tasty food, a man crying, or experiencing pleasure from watching boxing or a bull-fight are hereby

declared abnormal, at least in certain circles.

In physical medicine most normal or reference values are similarly based on a norm. Indeed, it is the only implied definition of normality in *Dorland's Medical Dictionary*.¹ The norm, in this case, seems to be the immediate post-pubertal physical state. This applies particularly to the systemic arterial blood pressure, body fat content, glucose tolerance, and plasma lipid profiles, all of which change with age. Although these changes are the rule, they are seldom considered to be normal. Indeed there are, in medicine, almost no unreservedly age-specific normal values for middle-aged and elderly persons.

The best

To a degree, the medical profession applies different definitions of normality at different times so as to render the normal state more desirable/optimal/favourable/advantageous than the abnormal state. Most of the above definitions of normality are therefore, to a greater or lesser extent, subservient to this definition.²⁻⁵

'Best', 'desirable', 'optimal', 'favourable' and 'advantageous' are, however, not internally defined concepts.^{5,10} They do not describe qualities such as 'green', 'solid', or 'silent', which are metric concepts that, in themselves, need no further clarification. 'Ideal', 'good', 'favourable' and 'advantageous' have no independent meaning. Weather that is, for instance, best for agriculture may not be best for the tourist industry. In fact nothing can be globally optimal, good, favourable, advantageous, or ideal. A car that is built for optimum economy, cannot simultaneously be best for speed or comfort.

If, therefore, an ideal determines the concept of normality, then its parameters need to be spelled out precisely. If the ideal is, for instance, the minimisation of the incidence of atherosclerosis, then the ideal *cannot* simultaneously be the prolongation of the useful life-span. Undernutrition and early death are by far the most effective means of eradicating atherosclerosis. If, more reasonably, the ideal is the maximisation of the useful life-span then, inevitably, the degenerative diseases of old age will become the commonest cause of death, (everyone must die eventually, and if they do so in old age, then the causes of death in old age must predominate). Instead of deploring such a trend, it would then become the stated goal of the public health services to have everyone die of the so-called degenerative diseases of old age.

The therapeutic definition

The McMaster Group⁶ provides a sixth definition of normality, which is not used in conversational English, but which has great clinical utility. Here 'normality' is that range of values (e.g. of arterial blood pressure or blood sugar) where therapy does more harm than good. Patients are therefore not labelled abnormal unless they are to be treated.

Like the previous definition, this definition relies on the concepts of 'good' and 'harm', and is therefore incomplete. It also means that all untreatable conditions (e.g. cancer and AIDS) are classed as normal, while everything that is currently treated is abnormal: fertility, the prepuce, and the orientation of the teeth in the mouth.

Discussion

It is clear that a disparate assembly of definitions is used to describe normality in medicine. The apparently arbitrary manner in which definitions are selected to evalu-

ate different phenomena effectively means that medicine operates without a definition: a certain group of phenomena are normal (by decree, as it were), and others are not. Actions based on such an arbitrary system are, of necessity, haphazard.

Of the definitions discussed above, only two are potentially acultural and objective: normal = usual, and normal = promoting useful longevity. For the medical profession the latter is probably preferable. Those conditions that shorten the useful life expectancy can then be termed 'diseases'. There are many caveats, however. The chief determinant of mortality is age. Indeed, age has few rivals as a predictor of impending death. Beginning at early puberty, the risk of death doubles approximately every 8 years. The risk of death is therefore more than thirty times greater at the age of 60 than it is at the age of 20 years. Misrepresentations are therefore sure to arise if comparisons are made across age groups. Thus, people who play bowls have a considerably higher mortality rate than pugilists have; physicians are less likely to survive until their next birthday than are their housemen; and teenagers 'cause' more strokes and heart attacks in their parents than toddlers do. Similarly, since obesity, grey hair, hypertension, hypercholesterolaemia, and non-insulin-dependent diabetes are characteristics primarily of middle- and old-age, they are all, *of necessity*, associated with substantially higher mortality rates than are their cultural norms. Only gender- and age-specific comparisons provide unbiased information on which the definition of disease can be based.

Since death cannot be eliminated, the health sciences can do no more than exchange one cause of death for another. Whether such an exchange is socially acceptable or not depends on whether the new cause of death kills people earlier, at the same age, or at a later age than did the previous cause of mortality. There are therefore no shortcuts when it comes to determining gender- and age-specific mortality rates. Smoking, for instance, predisposes to lung cancer. This, in itself, could well indicate a stupendous medical triumph, equal to the discovery of penicillin. If the average age of death from lung cancer is higher than the population average age of death, then smoking is clearly an elixir of life. The reason that smoking is in actual fact a health hazard, or 'disease', is that the average age of death is lower in smokers than in non-smokers.¹¹ The itemised list of causes of death will, in smokers and non-smokers alike, *always* total 100%. The composition of the list is therefore irrelevant to a discussion of what is most desirable.

Several of the factors which predispose to ischaemic heart disease, e.g. middle-age obesity and cholesterol levels, cause late rather than early all-cause death, compared with their *age-specific* norms.¹¹⁻¹⁴ They cannot therefore be termed 'diseases' in middle-aged people. Attempts at treating or preventing them are not in the patients' best interests (if 'best interests' involves prolonging the useful life-span).

Since disease is determined by strict gender- and age-specific criteria, diseases must perforce also be gender- and age-specific. Thus, beard growth is pathological at the age of 4 years, but entirely normal in 20-year-old men. A resting arterial blood pressure of 190/120 mmHg substantially shortens life expectancy at the age of 30 years, but has no effect on the mortality of 80-year-old men.¹⁵

Whatever definition of disease is used, the profession cannot operate without the 'normal = usual' definition of normality. This applies particularly to the items that constitute the familiar tables of 'normal values'.

These 'normal/usual' values are the source of great misgivings⁶ because, by definition, always exactly the same proportion (5%) of young asymptomatic adults fall outside the accepted range of normality. So no

matter what is measured — blood sugar, plasma sodium, the electrocardiographic PR-interval, height, weight, vital capacity — it appears (wrongly so) that 'all diseases have exactly the same frequency'.⁶ The problem is seemingly compounded by the fact that the probability of being declared normal decreases exponentially with the number of tests carried out on the individual ($P = 0,95^n$, where n is the number of independent tests carried out on the person).⁶ Thus, the probability of an asymptomatic young adult being extra-normal on the basis of a single test is, *by definition*, 5%. With 10 independent tests 40% (i.e. $1 - 0,95^{10}$) of young asymptomatic adults will have at least one extra-normal result. With 25 independent tests 72% (i.e. $1 - 0,95^{25}$) of the persons used to establish these normal values are found, *by definition*, to have extra-normal values. Clearly 'extra-normal', in this context, cannot mean 'diseased'.

When the features of a particular disease are described (irrespective of the definition of disease), it is generally considered superfluous to mention that patients with this condition invariably have 2 eyes, a nose, 2 arms, 10 fingers, and a pair of legs. These features, though truly characteristic of the disease (and which occur in over 99,999% of cases), are unhelpful because they are common to most other human diseases and also to health. Limitations with regard to time, memory, and printed space therefore dictate that the description of any given disease be confined to the list of rare features usually associated with that condition. It is for this purpose that the tables of common ($P = 0,95$) values exist. Values outside these ranges are merely rare; they are not intrinsically abnormal. On their own they are no more indicative of disease, or of health, than the possession of a common feature such as a pair of arms or a nose. In contrast to common features, however, a *combination* of rare features can economically pinpoint a diagnosis. A specific combination of five rare features will occur by chance only once in 3 200 000 ($P = 0,05^5$) asymptomatic people; this means that, for practical purposes, it never occurs in healthy individuals. The presence of this combination of features in a patient suspected of harbouring a disease known to be characterised by these features, is therefore 'diagnostic' of this disease. These rare features merely describe the condition, they do not cause it to be a disease. The condition qualifies as a disease because it increases the patient's gender- and age-specific mortality rate. Rare features economically describe any special group of people or situations: athletes in the immediate post-exercise period, high-altitude residents, chess-players, computer programmers, Shakespeare enthusiasts, etc.

It is therefore medically counter-productive to maintain lists of reference values which render extra-normality more common than $P < 0,05$ in any given gender or age group. Such lists can only lead to the proliferation of nonsense syndromes. These syndromes are no more than chance associations of more or less common features which describe nothing more useful than that the patient is, for instance, middle-aged. Consider, for instance the Smith syndrome: baldness, haemorrhoids, presbycusis, abdominal obesity, and midnight insomnia. Combinations of common features merely describe the species. Only exceptionally would a species characteristic usefully be classed a disease.

Conclusion

There is a clear need for a uniform, rational definition of normality in the medical sciences. The diversity of contradictory definitions currently in use cannot form the basis of a scientific endeavour. I have suggested what might usefully define a disease, and how 'normal/usual

values' aid the diagnostic industry. It could be argued that the age- and gender-specific mortality rate criterion for disease should be replaced by the corresponding morbidity rates. In other words, a quality-of-life instead of a length-of-life criterion should determine what is normal and what is not. However, how does one measure morbidity? How does one compare, for instance, the common cold with leukaemia? It is the risk of death that determines our subjective weighting of morbidity. Age- and gender-specific mortality rates, therefore, satisfy the desire for a morbidity-based definition of disease without being tautologous.

REFERENCES

1. Taylor EJ, ed. *Dorland's Illustrated Medical Dictionary*. 27th ed. Philadelphia, PA: WB Saunders, 1988: 1149.
2. Scadding JG. The principles of definition in medicine with special reference to chronic bronchitis and emphysema. *Lancet* 1959; **1**: 323-325.
3. Scadding JG. Diagnosis: the clinician and the computer. *Lancet* 1967; **2**: 877-882.
4. Campbell EJM, Scadding JG, Roberts RS. The concept of disease. *BMJ* 1979; **2**: 757-762.
5. Scadding JG. Health and disease: what can medicine do for philosophy? *J Med Ethics* 1988; **14**: 188-124.
6. Department of Clinical Epidemiology and Biostatistics, McMaster University Health Science Centre. How to read clinical journals: II. To learn about a diagnostic test. *Can Med Assoc J* 1981; **124**: 703-710.
7. Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc R Soc Lond (Biol)* 1979; **205**: 581-598.
8. Gombrich EH. The visual image. *Scientific American* 1972; **227**(3): 82-97.
9. Gould SJ. *Hen's Teeth and Horse's Toes*. Harmondsworth: Penguin, 1987: 177-198.
10. Lewontin RC. The shape of optimality. In: Dupré J, ed. *The Latest on the Best. Essays on Evolution and Optimality*. Cambridge, MA: MIT Press, 1987: 151-159.
11. Sorlie P, Gordon T, Kannel WB. Body build. The Framingham study. *JAMA* 1980; **243**: 1828-1831.
12. Gordon T, Kannel WB. Obesity and cardiovascular disease. *Clin Endocrinol Metab* 1976; **5**: 367-375.
13. Kozarevic D, McGee D, Vojvodic N, et al. Serum cholesterol and mortality: the Yugoslavia Cardiovascular Disease Study. *Am J Epidemiol* 1981; **114**: 21-28.
14. Kagan A, McGee DL, Yano K, Rhoads GG, Nomura A. Serum cholesterol and mortality in a Japanese-American population: the Honolulu Heart Program. *Am J Epidemiol* 1981; **114**: 11-20.
15. Holme I, Waaler HT. Five-year mortality in the city of Bergen, Norway, according to age, sex and blood pressure. *Acta Med Scand* 1976; **200**: 229-239.

REVIEW ARTICLE

Exercise and the gastro-intestinal tract

J. P. WRIGHT

Abstract Approximately 50% of athletes will develop gastro-intestinal symptoms at some stage in their careers. These range in severity from heartburn to gastro-intestinal bleeding. Fortunately symptoms are usually mild and inconvenient, but in certain individuals they can be incapacitating. It is important to exclude the more common gastro-intestinal conditions before diagnosing exercise-related syndromes. However, once such a diagnosis has been made, therapeutic options are limited.

The physiological role of the gastro-intestinal tract in fluid and energy replacement is increasingly being recognised. Without adequate replacements, performance may be limited. The volume of fluid ingested during endurance events needs to be limited to actual requirements; 500 ml/h is the average. Greater volume intake may be associated with overhydration and hyponatraemia. Glucose supplementation is essential for adequate performance in events of 2 - 3 hours' duration or longer. Studies of hyperosmolar carbohydrate solutions and their influence on energy and fluid emptying from the stomach suggest that higher carbohydrate concentration solutions than those often used by athletes may be advantageous.

S Afr Med J 1993; **83**: 50-52.

ANYONE who has had the pleasure of partaking in equestrian sports will know the effect of relatively mild exercise on the horse's gastro-intestinal system. In man, cardiovascular and respiratory limitations on performance and the value of cardiovascular training in improving performance in aerobic sports is well recognised. The role of the gastro-intestinal tract, both as a limiting and sustaining factor in aerobic exercises, is less well appreciated.

Gastro-intestinal symptoms

The spectrum of gastro-intestinal effects of exercise is wide (Table I), and ranges from the unlikely observation that swallowing decreases oxygen saturation in patients with chronic obstructive airways disease¹ and, by inference, in exercising athletes, to the more serious rectal bleeding that may be due to ischaemic bowel.²

The symptoms experienced by marathon runners are well known,³ with over 50% of runners at some stage of their careers reporting an urge to defaecate and just under 50% actually passing stools at least once while running (Table II). Abdominal cramps and anorexia are

TABLE I.
Gastro-intestinal effects of exercise

Swallowing
Gastro-oesophageal reflux
Gastric emptying
Gastric acid secretion
Peptic ulcers
Gastro-intestinal blood supply
Intestinal motility
Pancreatic function
Colonic function

Gastro-intestinal Clinic, Groote Schuur Hospital and
 Department of Medicine, University of Cape Town
 J. P. WRIGHT, M.B. CH.B., F.R.C.P., PH.D.

Accepted 7 Nov 1991.