Commentary: Origins and evolution of body mass index (BMI): continuing saga

Henry Blackburn* and David Jacobs Jr†

Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, MN, USA

*Corresponding author. Professor Emeritus, Division of Epidemiology and Community Health, University of Minnesota, 1300 S. Second St., Minneapolis, MN 55454. E-mail: Black002@umn.edu
†The authors contributed equally to this work.

We reflect upon Ancel Keys’ classic article, reprinted here, which dealt with a leitmotif of his long career: body mass, its composition, measurement, function and meaning for health, disease and survival. This preoccupation was...
Brozek in the Laboratory of Physiological Hygiene at other monumental study carried out earlier with Joseph here to validate the relative weight indices derive from an-

characteristics of risk and the prevention of common logical hygiene and public health. This led, in turn, to his development of concepts, beyond physiology, in physio-
garded as fixed ‘constitutional’ traits, were central to his anatomical and functional characteristics previously re-
position and Sheldonian body typing, and in numerous starvation experiment, of dramatic changes in body com-
believe that Keys' observations, in his Minnesota semi-

over height (m) squared.2

terize relative body weight, that is, the ratio of weight (kg) over height (m) squared.2

It was in the 1972 article, reproduced here partly because of its voluminous citations, that Ancel Keys gave Quetelet’s calculation its modern name, body mass index (BMI), along with evidence to support its current wide usage.

To begin, we particularly recommend Keys’ short intro-
tueto the characteristic clarity and par-
simony of his writing and for his lucid thinking about the measurement and meaning of body mass. In the text, he compared historic indices of relative weight for their independ-
ence of weight from height and for their representa-
tion of body fatness. The latter was based on correlation of an index with direct measure of body fat, either by the practical method of skin-fold thickness or the ‘platinum meter’ of laboratory-determined body fat. We then refer to the example of epidemiological design in which Keys tested the several indices across populations of men varying widely in body size and build.

We happen to know that Keys himself did all these analyses and wrote this article, while his co-authors, all long-term colleagues, provided the cohorts and made the field measurements. This effort prevailed among larger at-
tempts to relate obesity and body build to incident disease, death, or survival, which were among the core pursuits of Ancel Keys’ career.

Keys’ interest in and contributions to the detailed anthropometry involved in the indices tested here date to his earlier work on nutrition, body composition and performance in studies of semi-starvation and other stresses imposed among conscientious objectors to war and among army personnel during and after World War II.3 We believe that Keys’ observations, in his Minnesota semi-
starvation experiment, of dramatic changes in body composi-
tion and Sheldonian body typing, and in numerous anatomical and functional characteristics previously re-
garded as fixed ‘constitutional’ traits, were central to his development of concepts, beyond physiology, in physio-
logical hygiene and public health. This led, in turn, to his ideas and enterprise about the potential for modifying characteristics of risk and the prevention of common diseases.

The quantitative measures of body fatness Keys used here to validate the relative weight indices derive from an-
other monumental study carried out earlier with Joseph Brozek in the Laboratory of Physiological Hygiene at Minnesota and published in 1953 as ‘Body fat in adult man’.4 That study cross-validated various measures of body composition in fat and bone, ‘lean mass’, ‘fat-free’ mass and the plasma and tissue fluid or water ‘compartment’. One of us (H.B.) arrived on the scene in the Laboratory at the tail end of this extensive study and was promptly immersed, not only in the actual underwater weighing tank—followed by a seven-minute nitrogen washout to approximate lung volume for the body density calculation—but also in methods for estimating the body fluid compartment with heavy water or antipyrine or radioisotopes (all rapidly and evenly distributed through-
out body water and having predictable rates of disappear-
ance). Even more exotic measures, for example the volume of gastrointestinal gases, were considered at the time to refine the body density computation (e.g. colonic flatus aver-
ages 50 ml!).

The 1950s were a golden era of quantitative human biology in Keys’ laboratory housed in the bowels of the University of Minnesota football stadium. It was a time of challenge in its reductive and massive detail and, for some of us, overwhelming in its seeming reach for a ‘generalized field theory’ of human biology. In any case, body compart-
ments and their functionality were at the centre of Keys’ physiological hygiene, which, in turn, was a product of his multidisciplinary investigative strategy. Soon this approach would be applied to questions surrounding the ‘diet-heart’ hypothesis by comparisons within and among cohorts of widely varying eating patterns and similar traditional occupa-
tions, in the Seven Countries Study.5

Insurance industry relative and ideal weight tables

Keys’ first analysis deals with the simplest expression of relative weight, that used by life insurance companies in describing the weight of an individual as a percentage of the average weight of persons of the same height, age and sex in a given population. These average values came to be considered ‘normal’ weight, then ‘standard’ weight, which soon morphed into ‘ideal’ weight based on actuarial investigations of mortality in 1912 and 1959.6 These tables and their implications about excess risk of overweight not only engaged but irritated Keys throughout his professional life, during which he often criticized them, based on his finding that people of standard weight at different ages differed greatly in body fat content. He maintained that validated, quanti-
tative height-weight relationships were essential to any ‘scientific’ explorations of body mass, its meaning and its changes over time. The article reprinted here was his attempt to provide this essential element.
The ponderal index

Dealing thus abruptly with ‘industry-hyped’ tables, Keys then asked how body mass, which is equivalent to a volume and thus three-dimensional, could be standardized at all in reference to a singular, linear dimension of height? He reasoned that the body would not necessarily have the same form at different heights, as was assumed by a formula in vogue then for adults (and still for children) as the ponderal index, weight (kg) over height cubed (m). The inverse of this index was championed by the Sheldonian school of body habitus classification, which Keys labelled in this article as ‘bizarre’ and having ‘the unhappy feature of being inversely proportional to weight at any given height’.1

Body mass index

Keys’ accurate provenance for the origin of a particular relative weight index takes us back to mid-19th century and the story of Adolphe Quetelet. At age 23, the remarkable Belgian mathematician went to study in Paris under Poisson, Laplace and Fourier and returned to found the Royal Astronomical Observatory, then to develop and promote the sciences of statistics and of anthropometry. All was part of his grand scheme to measure and characterize ‘l’homme moyen’—average man—by the mean values of measured variables having a normal distribution!2 (Quetelet’s quantitative social concepts became so popular that he was called on to counsel Prince Albert in London on politics and President Garfield in Washington on the US census, and to philosophize in a long correspondence with Goethe).3

Keys proposed in this 1972 article that the index devised by Quetelet, the ratio of weight over height squared, now be called the ‘body mass index’ (BMI), the term that persists today, along with its predominant usage in the quantitative study of body mass and obesity, in health and disease.

Indices against a reference standard; seeking variability

Keys studied how the several historical indices of relative weight stacked up against direct measures of body fatness, maintaining that:

In both medical and popular uses of relative weight data the interest, conscious or unconscious, is on the implication for body fatness. It is of interest therefore to examine the relationships between the various indices of relative weight and completely independent measures related much more directly to the body fat mass. This he proceeded to do in populations widely contrasting in size and weight, setting the tone as follows:

Other things being equal, it is agreed…that the best relative weight index is the one that shows the least correlation with body height and the highest correlation with independent measures of body fatness.1

Keys began in the early 1950s to seek out wide variations among populations as key to the search for causes in chronic disease epidemiology, which had initially preoccupied itself with individual risk within homogeneous, affluent, White populations (e.g. the Framingham Study, the British Doctors Study and Keys’ own vastly underpowered early study among such a cohort of men in Minnesota).

Keys’ comparative findings

Keys found all of the indices of relative weight highly correlated with thickness of the subcutaneous fat layer, with only small differences between them. The correlations showed trivial differences using the transformed values for skin folds, which have highly skewed distributions. With body fatness as the reference, the BMI proved to be ‘if not fully satisfactory, at least as good as any other relative weight index as an indicator of relative obesity’.1

Keys concluded from his measures in groups of men, tall and short, heavy and light, that:

By the criteria of correlation with height (lowest is best) and two measures of body fatness (highest is best), ponderal index is the poorest of the relative weight indices studied. The ratio of weight to height squared, the body mass index, is slightly better in these respects than the simple ratio of weight to height. The body mass index seems preferable over other indices of relative weight on these grounds as well as on the simplicity of the calculation and, in contrast to percentage of average weight, the applicability to all populations at all times.1

Keys’ opinions, attitudes and aesthetic

In Keys’ lifelong preoccupation with body composition and fatness he maintained that overweight was not necessarily ‘over-fat’. He also was aware of the potential for an obesity epidemic in affluent society as early as 1949, and in 1950 proposed dietary change rather than increased physical activity as the best approach to weight reduction.7,8 In 1952 he criticized Sheldon’s somatotyping of body habitus as ‘a devious and inefficient route’ for estimation of body fatness.9

For many years prior to this 1972 report, Keys had proclaimed the inadequacy of relative weight as a measure of fatness, pointing out the misinterpretations in muscular people and those with metabolic abnormalities. (He, for example, was short in stature and highly ‘muscle’d’ and dense.) This led him to compare contrasting populations in the USA and northern Europe, the Mediterranean basin
and Asia for the development of coronary heart disease (CHD) and death as a function of obesity and of habitual diet. In this larger search for the implications of overweight and obesity, he found in the long-term prospective Seven Countries Study that there was a positive association of overweight with angina pectoris but not with myocardial infarction or coronary death and that the relationship to total mortality was U-shaped. In the final paper of his life he found that weight gain during middle life was greater in men who survived to advanced age compared with those who had died, even among men who had never smoked or had given up smoking.

Throughout his writings, Keys’ aesthetic view of obesity ranged from negative to devastating, as when he called obesity ‘disgusting as well as a hazard to health . . . ethically repugnant, uncomfortable and impedes motion . . . hard on clothes and furniture’. The very fat are ‘clumsy and prone to accidents’. He maintained that obesity was ‘ugly but does not itself cause CHD’, and was not necessarily dangerous for mortality risk in the average working man in the traditional populations we studied in mid century. He showed that modest ‘overweight’ may have advantages, particularly in the elderly, and he established that maintenance of low body fatness facilitates a favourable cardiovascular risk profile. He fully understood the dangers of mechanism of life’s tasks and of overabundant food and warned early of the impending obesity epidemic.15 He arrived at that conviction by careful and thoughtful analysis, weighing his findings with his and others’ prejudices. We also note that he seemed surprised that BMI worked so well, and perhaps he did not fully understand why. All things considered, he recommended it to estimate fatness because it is right so often and is so simple.

The simplicity of BMI, and the similar scaling of adipose tissue compartments (defined by magnetic resonance or X-ray absorptiometry) to height, have continued to surprise researchers who investigate specific fat depots. Over the years since Keys’ 1972 summation, many have pondered the limitations of BMI, questioned its independence of stature, particularly at changing phases of the life cycle, and proposed coefficients for a ‘new BMI’. An Oxford mathematician, Nick Trefethen (like Quetelet, neither physiologist nor epidemiologist!), has recently proposed that the small over-classification of BMI among tall persons and under-rating among the short would be corrected by a new ratio: weight times 1.3 (i.e. the square root of average height, or 1.69 m) over height to the 2.5 power, and reflect better how weight relates to height in ‘most healthy adults’! Where are Keys and colleagues to validate such apparently clever ideas?

In fact, huge databases with BMI measurements from long-term prospective studies are now available on which such ‘new BMI’ formulations might be tested against the standard. Nonetheless, standard BMI has survived as a robust, useful and surprisingly accurate measure of fatness in ‘healthy’ adults.

**Conflict of interest:** None declared.
The axiom that body weight increases with height stimulated Ancel Keys and colleagues to address the public health quandary of the appropriate body weight for health and longevity. Awareness of the limitations of the standard height-weight tables from the life insurance industry established the need for a simple, physical indicator of body size that differentiates body fatness levels with increasing weight of adults in the assessment of risk of mortality. Keys et al. cleverly integrated the resources of physical anthropology, nutrition and physiology and demonstrated that body weight divided by standing height squared (Wt/Ht²; kg/m²) or the body mass index (BMI), compared with other weight and height ratios, was a valid indicator of adiposity because it increased with greater levels of body fatness among groups of men. This critical finding evolved to the development of specific ranges of BMI to categorize underweight (<18.5 kg/m²), normal (18.5–24.9 kg/m²), overweight (25–29.9 kg/m²) and obesity (>30 kg/m²) with gradations (class I, 30–34.9; class II, 35–39.9; and class III >40 kg/m², respectively). Thus BMI was designated a measure of weight status, and serves as a prominent variable in population research to elucidate relationships between adiposity and metabolic variables associated with increased risk of cardiovascular disease, stroke, type 2 diabetes, cancer and mortality.

Widespread use of BMI as a measure of adiposity and a predictor of health risk requires an understanding of its limitations for an individual. Despite consensus that BMI accounts for appreciable variance (60–70%) in measured fatness in groups of adults, it is an unreliable indicator of the body composition of an individual. Correlation coefficients relating BMI to fatness and fat-free mass are similar (r = 0.7 to 0.8) and indicate its non-specificity in prediction of body composition for an individual. BMI does not adequately distinguish fat from muscle and bone, so individuals with the same BMI can have vastly different amounts of muscle and body fat. Also BMI does not account for age, gender, ethnicity and physical fitness in...