Menarche and Fatness: Reexamination of the Critical Body Composition Hypothesis

In a series of articles published during the last few years (*1–8*), Frisch and her colleagues have developed the hypothesis that menarche requires a critical level of fat stored in the body. They reason that pregnancy and lactation impose a great caloric drain; if fat reserves are inadequate to meet this demand, then fecundity is impaired.

This theory, with its implications for fertility, is attractive because it is both simple and consistent with the observed and well-documented delay of menarche among severely malnourished girls and loss of menstrual cycles during famines. It has long been thought that nutrition may play a role in reproduction, perhaps affecting conception, fetal mortality, the health of the newborn, or the length of postpartum insusceptibility. There is reliable evidence that in some historical populations which do not appear to have practiced birth control age-specific marital fertility rates were only half the observed maximum. If, as Frisch asserts, poor nutrition delays menarche, lengthens the periods of adolescent sterility and postpartum amenorrhea, and lowers fecundability, then the very low fertility observed in these historical populations could perhaps be easily explained (*3*).

A close examination of the evidence Frisch and her colleagues offer suggests, however, that their reasoning is flawed either by statistical error or by selective interpretation of the data. These findings do not invalidate the hypothesis, but establish clearly that more careful testing is needed.

In support of the hypothesis they offer five pieces of evidence:

1) Sets of weight-for-height standards which have proved to be useful in predicting a minimum weight (for a given height) necessary for the initiation of menses in adolescents or return to menses in anoretic patients are based upon an index of fatness (*6*).

2) The coefficient of variation of weight at menarche is significantly higher than the coefficient of variation of the ratio of total body water to total weight, a ratio which they estimate from height and weight and which in turn is their index of fatness (*8, 9*). They emphasize that this finding has biological significance because it identifies fatness as the critical factor governing menarche.

3) Regression equations which predict age at menarche from height and weight yield significantly better estimates when girls are stratified by the index of fatness. A separate regression equation is provided for girls in each quartile of fatness (*4*).

4) A higher proportion of girls remain in the same quartile of fatness than in the same quartile of weight or body water from the time of initiation of the adolescent growth spurt to menarche (*8*, p. 477).

5) Historical data published by Quetelet conform to the hypothesis and establish a historical constancy of average weight at menarche of 46 kg (*5*).

Before we proceed further, a brief explanation of Frisch's methodology is essential:

Total body weight (*WT*) is regarded as the sum of lean body weight and fat. Body water (*TW*) is considered to be a constant 72 percent of lean body weight and is estimated from observations on height (*HT*) and weight by means of the following regression equation derived by Mellits and Cheek (*10*):

$$\hat{TW} = -10.313 + 0.252\ WT\ (\text{kg}) + 0.154\ HT\ (\text{cm}) \quad (1)$$

where the hat over *TW* indicates that it is estimated and not observed. Since $WT = TW/0.72 + \text{fat}$, the ratio TW/WT is an inverse index of fatness, which from Eq. 1 can be estimated as a function of height and weight alone:

$$\hat{TW}/WT = -10.313/WT + 0.252 + 0.154\ HT/WT \quad (2)$$

At several points during the following discussion reanalysis of the longitudinal data used by Frisch would be desirable. Unfortunately, the only data available in published form are those of the Berkeley Guidance Study (BGS) (*11*), which is the source of roughly a third of the sample employed by Frisch.

The height and weight standards. Using observations on height and weight at menarche obtained from the BGS and from studies conducted by the Child Research Council and the Harvard School of Public Health, Frisch and McArthur estimated the ratio TW/WT for each of 181 subjects. The distributions of this ratio at menarche and at age 18 were formed and centiles were calculated. For example, 10 percent of the girls at menarche were found to have a $\hat{TW}/WT$ value higher than .598. Hence a line representing the tenth centile can be calculated from Eq. 2 as

$$.598WT = -10.313 + 0.252WT + 0.154HT \quad (3)$$

Using such loci, along which relative fatness is supposedly constant, Frisch and McArthur constructed weight-for-height standards for menarche and for restoration of the menses in amenorrheic women (*6*). Frisch has suggested that the tenth centile locus represents a critical body composition. Girls below the tenth centile do not have sufficient stored fat and will not menstruate until they gain enough weight to push them above the tenth centile. It was my intention to reproduce the figures presented in (*6*) by Frisch and McArthur, but as they refused me permission to do so, and as Frisch also refused me the data used to construct them, Figs. 1 and 2 here are derived from Eq. 2. They are not as rich in detail as they would otherwise be but they portray the general notion. The tenth centile in Fig. 1 represents a body composition of 17 percent fat; the tenth centile in Fig. 2 corresponds to a higher level of 22 percent fat necessary for restoration and maintenance of menses. Because of their choice of the tenth centile Frisch and McArthur expect a 10 percent error in classification. It should be noted that the locus is critical in one direction only; being above the locus is necessary but not sufficient for initiation or return of menses.

It should be noted also that the fact that the weight and height standards have been found to be useful in some predictions or diagnoses does not validate the underlying theory. It is possible to formulate alternative theories which result in operationally equivalent standards, in the sense that they would lead to the same clinical recommendations (*12*).

Since the centiles in Fig. 1 are constructed from Eq. 3, they are in effect centiles of observed weights and heights, not of fatness, which is merely estimated from them according to a formula. Hence a better set of standards could be constructed without reference to any underlying theory whatsoever. A straightforward data analysis of weight and height at menarche would yield more internally consistent loci. For example, a simple regression of weight at menarche on height at menarche would allow construction of centiles whose slopes reflect the structure of the data and not a theoretical construct. For the 65 girls in the BGS that regression is estimated to be

$$WT = .622HT - 49.513 \quad (4)$$

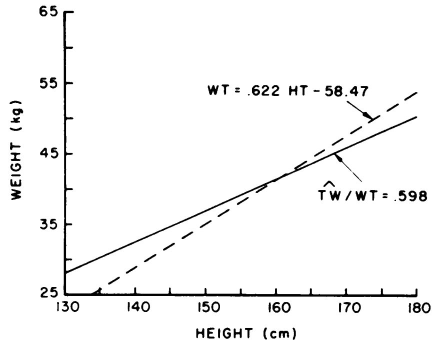
Standards can be constructed by simply varying the constant in Eq. 4. One locus is plotted as a dashed line in Figs. 1 and 2 (*13*). A similar analysis could be performed on the heights and weights of

previously amenorrheic patients at the time of return to menses to yield more internally consistent standards for Fig. 2.

The coefficient of variation. In her search for the factor which controls menarche, Frisch has examined what women have in common at the point in their lives when menarche occurs. They do not appear to have a common age, weight, or height, but they do have a common (estimated) ratio of total water to total body weight as measured by the coefficient of variation (CV = standard deviation/mean).

Billewicz *et al.* (*14*) have shown that this reduction in CV is purely the result of a mathematical identity. The standard deviation and mean, and therefore the CV of the fatness index $\hat{TW}/WT$, are determined solely by the observations on height and weight (*15*). Whether the CV of the estimated ratio is lower or higher than the CV of WT therefore could have no biological significance (*16*). Suppose, however, that direct observations on TW were available and that the reduction in CV were found to be real. Is the CV a particularly revealing statistic in this context? It is attractive because it is dimensionless; it would have the same value whether TW/WT were measured as liters per kilogram or gallons per pound. Unfortunately, a large part of its denominator (WT) is accumulated prior to puberty; hence the statistic is in effect deflated. For example, if all girls wore platform shoes 30 cm high, then the CV of total height (with shoes) at menarche for girls in the BGS sample would be 15 percent smaller than that of height without shoes. Likewise, body temperature at menarche would have a lower CV if measured in degrees Kelvin than in degrees Celsius. This point is particularly relevant when discussing the estimated ratio TW/WT, since it is bounded away from zero at every age; its smallest value occurs among short, fat girls; a reasonable lower bound of .42 might occur for a girl whose height and weight are 110 cm and 40 kg respectively. If one regarded the natural zero of the ratio as .42, then the coefficient of variation would be 4.2 times as large as that reported by Frisch (*17*); but since the ratio has no natural zero, comparisons employing its mean and coefficient of variation are totally without meaning.

Predicting age at menarche. Frisch has devised a procedure for predicting



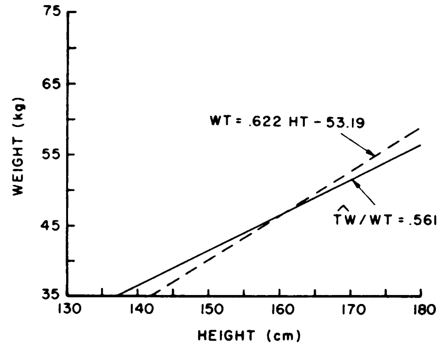


Fig. 1 (left). The minimum weight necessary at a particular height for onset of menses is represented by the solid line for the Frisch and McArthur standards and by a dashed line for the standards obtained by regression of weight at menarche on height at menarche. Fig. 2 (right). The minimum weight necessary at a particular height for the resumption of menses is indicated by the solid line for the Frisch and McArthur standards and by a dashed line for the standards obtained by regressing weight at menarche on height at menarche. The solid lines (Frisch and McArthur standards) were derived by substituting, respectively, .598 and .561 into Eq. 2.

Table 1. Sum of squared residuals (SSR) obtained from the regression of age at menarche on a constant and either height or weight. The F ratio for segmentation is given in the last row for each age. The letter beside each SSR indicates whether the regressor was height (H) or weight (W).

Age	Quartile	N	Total (over all subjects)	$\hat{TW}/WT$	HT	WT	WT/HT	$\hat{TW}$
9	1	16		13.9885 *H*	14.6880 *W*	11.9428 *W*	17.4529 *W*	15.0482 *W*
	2	16		6.5913 *W*	10.1095 *W*	18.5082 *W*	8.4101 *H*	10.2741 *H*
	3	16		7.8654 *W*	10.4110 *H*	7.7604 *W*	8.7201 *W*	9.8381 *H*
	4	17		20.4596 *W*	16.7642 *W*	11.8453 *H*	11.8453 *H*	11.4228 *H*
	Total	65	55.8819 *W*	48.9048	51.9727	50.0567	46.4284	46.5832
	F(6,57)			1.3553	0.7146	1.1055	1.9343	1.8963
10	1	16		9.7558 *H*	11.8363 *W*	13.1568 *W*	14.3379 *W*	12.6949 *W*
	2	16		8.6679 *H*	14.0664 *W*	14.8433 *W*	15.0483 *W*	12.2149 *W*
	3	16		7.4138 *H*	6.5297 *H*	6.7116 *W*	6.3783 *H*	12.4351 *W*
	4	17		18.3366 *W*	16.0149 *W*	9.9149 *H*	9.9149 *H*	11.3358 *H*
	Total	65	51.5319 *W*	44.1741	48.4473	44.6266	45.6794	48.6807
	F(6,57)			1.5824	0.6049	1.4700	1.2172	0.5564
11	1	16		9.4564 *H*	8.9511 *W*	11.6178 *W*	12.4922 *W*	12.0031 *W*
	2	15		3.6261 *W*	8.7422 *H*	8.0377 *H*	7.9780 *H*	8.0955 *H*
	3	15		7.7981 *H*	3.3624 *H*	5.9297 *H*	5.3204 *H*	6.0750 *H*
	4	16		12.9300 *W*	14.0243 *W*	7.8625 *H*	7.4025 *W*	7.0434 *W*
	Total	62	44.0385 *W*	33.8106	35.0800	33.4477	33.1931	33.2170
	F(6,54)			2.7226	2.2984	2.8497	2.9406	2.9320
12	1,2	25		8.9137 *H*	15.2453 *W*	14.9420 *W*	15.9764 *W*	11.5099 *W*
	3,4	25		14.1510 *W*	10.1941 *W*	6.2397 *W*	6.0260 *W*	11.3817 *H*
	Total	50	27.3829 *W*	23.0647	25.4394	21.1817	22.0024	22.8916
	F(2,46)			4.3061	1.7571	6.7335	5.6245	4.5126
13	1,2	13		2.2441 *H*	4.7094 *W*	6.3850 *W*	4.9872 *W*	6.1912 *W*
	3,4	14		8.5582 *W*	3.7797 *H*	3.7453 *H*	5.5138 *H*	3.6977 *H*
	Total	27	12.6451 *W*	10.8023	8.4891	10.1303	10.5010	9.8889
	F(2,23)			1.9618	5.6300	2.8548	2.3481	3.2052

age at menarche which is being widely used (*4*, p. 384). From the same body of data used to construct the height and weight standards, quartiles of $\hat{TW}/WT$ were constructed at each age from 9 to 13 (*18*). Each girl who had not yet reached menarche was classified into one of four quartiles at each age. Within each quartile, at each age, age at menarche was regressed as a function of height or weight. The equation that yielded the highest significant R^2 was chosen (*19*). The same procedure was repeated for height, weight, and weight-for-height quartiles. Frisch found that segmentation into $\hat{TW}/WT$ quartiles gave higher significant R^2 than when all subjects were combined at each age and that segmentation into height or weight or weight-for-height quartiles "gave either worse standard errors of the estimates" (lower R^2) than classification by $\hat{TW}/WT$ quartiles, "or insignificant results" (*4*, p. 386).

These findings are indeed interesting, but Frisch presented no test of whether the segmentation itself was necessary. Such a test can be formulated rather easily, and can be applied to the BGS subjects.

If the same regressors had been used in the regression equation for each quartile, then the test would be a straightforward F test (*20*). However, in some quartiles Frisch found that weight was more significant and in others that height was more significant; hence the assumptions of the standard F test are violated. Nevertheless, because we are interested in the significance of segmentation and not in the actual functional form (for which the only tests are complicated asymptotic ones), we can form an F test which is biased in favor of rejecting the null hypothesis that pooling is not a significant restriction. For each quartile we pick the lower of the SSR's (sum of the squared residuals) from the regression on weight or height. We next compare the sum of the SSR's over all quartiles at each age with the lower of the pooled SSR's resulting from a regression on either height or weight (*21*) in a conventional F ratio. That this test is biased in favor of rejecting the null hypothesis and accepting segmentation is obvious: the numerator of the F ratio must be biased upward, and the denominator downward. No conclusion can be drawn if the F ratio is significant; but if it is not, one can conclude that segmentation is not warranted. Results for the BGS data are presented in Table 1. It can be seen that no F ratio is even remotely significant when the data are segmented by $\hat{TW}/WT$ except at ages 11 and 12, and even at these ages segmentation by weight or *WT/HT* quartiles produces a higher F ratio. We conclude that either no predictive power is lost when the segments are pooled or, if it is, segmentations other than $\hat{TW}/WT$ perform equally well or better.

Table 2. Percent of BGS subjects who remained in the same quartile from age 8 through age at menarche.

Height	46.2
Weight	40.0
WT/HT	44.6
$\hat{TW}$	41.5
$\hat{TW}/WT$	43.1

Consistency of classification into quartiles of $\hat{TW}/WT$. Frisch says that "of the 169 girls who could be followed from [the initiation of the adolescent growth spurt] to menarche, 138 (82%) remained in the same quartiles of total water/body weight from initiation of the spurt to menarche, compared to only 79 (47%) remaining in the same quartiles of weight from initiation to menarche, and 66 (39%) remaining in the same quartiles of total water from initiation to menarche" (*8*, p. 477). She does not mention the percentage who stay in the same quartile of height or *WT/HT*. This question has been examined for the 65 girls in the BGS; unfortunately, insufficient detail was provided by Frisch to allow the unambiguous determination of the timing of the adolescent growth spurt. Some notion of the answer can be gained by following girls from age 8 to menarche. The percentages who remain in the same quartile of height, weight, *WT/HT*, $\hat{TW}$, and $\hat{TW}/WT$ from age 8 to menarche are given in Table 2. A higher percentage do indeed remain in the same $\hat{TW}/WT$ than remain in the same weight quartiles, but a still higher percentage remain in the same height and *WT/HT* quartiles. It would appear that the superiority of the performance of $\hat{TW}/WT$ is due to a constrained, selected comparison; a full comparison based on all 169 subjects would be in order. It is unclear, however, how this finding, even if correct, implicates fatness as a determinant of menarche.

The historical evidence. For historical evidence Frisch places great emphasis on the Belgian data collected by Quetelet. For example, when talking about the interval from the initiation of the growth spurt to menarche she states that "Belgian girls of over a century ago apparently grew more slowly *during* the adolescent growth spurt as well as *before* the spurt, so that this interval to menarche was about 4.5 years (Quetelet, 1869)" (*3*, p. 18). One might get the impression from the attribution to Quetelet that he discussed age at menarche, but as Frisch states elsewhere (*5*, p. 448), he did not. Instead, he talks about age at puberty (apparently the time of appearance of secondary sex characteristics), which he places at about 12 years. Frisch argues that the Quetelet data on weight by age for girls are consistent with an age at menarche between 16 and 17 years, and that this finding in turn provides evidence for historical constancy of an average weight at menarche of 46 kg, an important piece of corroborative evidence for her theory.

A closer examination of the Quetelet data (*22*) reveals, however, that her argument is not unambiguously sound. We have available four pieces of evidence: height and weight profiles for boys and girls by age. Regardless of any quibbles about critical weight or body composition, one is certain that menarche follows rather closely the peak adolescent growth spurt, as Frisch herself has noted (*3*, p. 18). Single-year velocity curves can be derived from the height-by-age or weight-by-age profiles by differencing. Three of these curves have well-defined peaks: male height, at age 15; female height, at 13; and male weight, at 16. These together indicate that menarche should have occurred no later than age 15, since the interval from the peak spurt to menarche is approximately 1 year (*23*), and the sex difference between peaks averages about 2 years (*24*). The velocity curve for female weight does not have a well-defined peak; instead two peaks, one at age 12 and the other at age 17, appear. It should be noted that the height profiles are more likely to be correct, because height was measured with the subjects shoeless, but weight was measured when they were clothed. To avoid the upward bias in weights, Quetelet subtracted, at each age, 1/18 and 1/24 of total weight for males and females respectively. With these considerations in mind, it is difficult to accept Frisch's argument that menarche occurred as late as 16 to 17 years.

Frisch has recently provided a long catalog of historical citations which support her theory of a direct link between nutrition and the biological capacity to reproduce (*25*). While this summary is impressive for its sheer magnitude, the fact that 18th- and 19th-century writers believed that link existed does not prove that it did; many of the same writers believed in other theories which today can

be rejected conclusively. What is needed is direct human corroboration, not mounds of historical evidence. Her work has inspired at least two direct tests. Bongaarts and Delgado, when analyzing the effects of nutritional status on fertility in rural Guatemala, found no effect on fecundability (monthly probability of conception) and a statistically significant but demographically unimportant lengthening of the period of postpartum amenorrhea by 1.6 months among women of low nutritional status (*26*). Another study conducted among rural women in Bangladesh found that the length of postpartum infertility was unrelated to maternal nutritional status (*27*) and that there seems to be no threshold effect of a minimum weight-for-height necessary for the resumption of menses following a birth. Neither of these findings supports the Frisch hypothesis and neither was mentioned by her.

Summary. The theory that fatness is a critical determinant of menarche or fecundity is plausible and merits investigation, but the evidence thus far produced for it is weak indeed. It remains to be seen whether the theory is correct and, if so, whether fatness (or nutrition) is a demographically important determinant of fecundity and thereby fertility. For example, if the period of postpartum amenorrhea were lengthened by 3 months by poor nutrition, the result would be statistically and biologically significant, but demographically unimportant. It is also important to discover whether there is a direct physiological link as Frisch proposes; nutrition might indeed affect fecundity but only indirectly through loss of libido.

JAMES TRUSSELL

Office of Population Research, Princeton University, Princeton, New Jersey 08540

References and Notes

1. R. E. Frisch, in *Control of Onset of Puberty*, M. Grumbach, G. Grave, F. Mayer, Eds. (Wiley, New York, 1974), pp. 403–423.
2. ——, in *Biosocial Interrelations in Population Adaptation*, E. Watts *et al.*, Eds. (Mouton, Paris, 1976), pp. 319–352.
3. ——, *Soc. Biol* **22**, 17 (1975).
4. ——, *Pediatrics* **53**, 384 (1974).
5. ——, *ibid.* **50**, 445 (1972).
6. —— and J. W. McArthur, *Science* **185**, 949 (1974).
7. R. E. Frisch and R. Revelle, *Arch. Dis. Child.* **46**, 695 (1971).
8. ——, S. Cook, *Hum. Biol.* **45**, 469 (1973).
9. In support of the hypothesis of metabolic mass as a trigger, Frisch states, "It is especially significant that the variability of TW [body water] as a percentage of body weight at menarche is 55 percent less than that of weight at menarche" (*1*, p. 415).
10. E. D. Mellits and D. B. Cheek, *Monogr. Soc. Res. Child Dev.* **35**, 12 (1970).
11. R. D. Tuddenham and M. M. Snyder, *Univ. Calif. Publ. Child Dev.* **1**, 183 (1954). Two subjects whose weights at menarche exceeded 70 kg were eliminated from the group so that the remaining 65 would be identical to those included in Frisch's combined data set.
12. It is not even necessary that an alternative theory be plausible in order for it to yield reasonable standards. For example, it could be hypothesized that there is a critical minimum intelligence necessary for the onset and maintenance of menses, on the grounds that women who do not meet this minimum would not be able to cope adequately with pregnancy, hence their fecundity is suppressed so that survival is ensured. I recruited six of my female colleagues and regressed their IQ's on their weights and heights to find
$$\hat{IQ} = 234.63 + 2.00WT - 1.25HT$$
Better standards than those offered by Frisch could be constructed from the reexpression of the previous equation as
$$WT = \text{constant} + 0.625HT$$
where the constant equals 0.5 ($\hat{IQ}$ − 234.63). For example, the locus $WT = -53.625 + .625HT$, along which $\hat{IQ} = 127.38$, if plotted as a dashed line in Frisch and McArthur's figure 2 (*6*) would discriminate perfectly between subjects who were menstruating and those who were not. These standards, like those of Frisch and McArthur, are weight-for-height standards. A woman of a given height increases her estimated IQ or fatness by gaining weight. Her actual IQ or fatness is irrelevant to the use of the standards.
13. One may regard Eq. 4 as a theory, to wit, there is a linear relation between weight and height at the time of menarche. It is certainly more plausible than the IQ theory (*12*), though they yield equivalent standards, since the slopes (.62) are equal.
14. W. Z. Billewicz, H. M. Fellowes, C. A. Hytten, *Ann. Hum. Biol.* **3**, 51 (1976).
15. For any equation of the form $y = a + bx_1 + cx_2$, it is always true that $E(y) = a + bE(x_1) + cE(x_2)$ and that $Var(y) = b^2Var(x_1) + c^2Var(x_2) + 2bc\, Cov(x_1x_2)$. Billewicz *et al.* (*14*) point out that one finding—that the coefficient of variation of total water is lower than that of weight—merely recaptures these two mathematical identities. Likewise the *CV* of $\hat{TW}/WT$ is determined by Eq. 2: $CV(\hat{TW}/WT) = \sqrt{a}/b$, where $a = (10.313)^2 \text{ var}(1/WT) + (.154)^2 \text{ var}(HT/WT) - 2(.154)(10.313) \text{ cov}(1/WT, HT/WT)$, and $b = -10.313 \text{ E}(1/WT) + .252. + .154 \text{ E}(HT/WT)$.
16. In the BGS sample the *CV* of $\hat{IQ}$ at menarche is 39 percent lower than the *CV* of weight at menarche—a fact which likewise has utterly no biological significance.
17. The 181 subjects had a mean ratio $\hat{TW}/WT$ of .551 (*8*, p. 472). If the natural zero is .42, then the coefficient of variation would be raised by a multiplicative factor of .551/(.551 −.42) =4.2.
18. Frisch defined age from "birthdate anniversary to next birthdate anniversary, e.g. age 9 is from the ninth birthday up to the tenth birthday." Therefore, the observations upon which regressions were based should have been centered on the midpoints of each age group. The observations on the BGS subjects are so centered, but the calculations have been made centering on exact age as well; the numbers change slightly, but the qualitative results are the same.
Frisch's quartiles of weight, height, and weight-for-height were determined from *national* samples; the quartiles of $\hat{TW}$ and $\hat{TW}/WT$ were, however, determined from her own sample. If one wishes to compare the performance of variables, it is more appropriate to use quartiles determined from the same sample. Since there are no national samples from which to compute $\hat{TW}$ and $\hat{TW}/WT$, all quartiles used here were determined from the data in the BGS sample. It is not possible to determine weight-for-height quartiles from such a small sample, however. Hence a ponderal index WT/HT was substituted; such a substitution does not alter the argument that segmentation was not necessary.
19. While this criterion may be used, it is more common, when one's aim is merely to predict, to choose the regression with highest R^2 rather than the one with highest significant R^2. If height and weight had been used as regressors in each quartile regression, then the predictions could not have been worse, as measured by the squared error loss function explicit in least squares regression. *F* tests of pooling all observations for each age with both weight and height as regressors show that only at age 12 can the null hypothesis (that the observations can be pooled) be rejected at the .05 level, and even for age 12 pooling produces higher *F* ratios when the data are segmented by WT/HT or by weight than when they are segmented by $\hat{TW}/WT$.
20. To test whether segmentation by quartiles is significant one runs five regressions for each age, one for each quartile and one for all subjects combined. The *F* ratio is formed as follows: (i) the *SSR*'s for each of the quartiles are summed; denote this sum *SSRu*, for unrestricted; (ii) the *SSR* for the single regression on the pooled observations is denoted *SSRr*, for restricted; and (iii) the quantity
$$F = \frac{(SSRr - SSRu)/3k}{SSRu/n - 4k}$$
is distributed as $F\ (3k,\ n - 4k)$, where n is the number of women at each age and k is the number of regressors (including the constant) in each regression.
21. One might wonder why the smaller of the two *SSR*'s resulting from the pooled regressions is used. The larger could also be used, but the test then becomes more biased in favor of rejecting H_0. We seek the better of the two biased tests. Even if the larger of the two *SSR*'s had been used, only at age 10 would pooling have produced the highest *F* ratio when the data are segmented by TW/WT.
22. M. A. Quetelet, in *Comparative Statistics in the 19th Century*, R. Wall, Ed. (Gregg, Farnborough, England, 1973), pp. 58–67. The data on height are taken from the first table in the chapter on height (p. 58) and the data on weight from the first table in the chapter on weight (p. 64). There are also data on height in the weight table; the peak spurts for both boys and girls occur at age 14. Quetelet apparently preferred the height data included in the chapter on height.
23. A review of the literature on this topic is found in J. Trussell and R. Steckel, *J. Interdiscip. Hist.* **8**, 477 (1978).
24. P. B. Eveleth and J. M. Tanner, *Worldwide Variation in Human Growth* (Cambridge Univ. Press, Cambridge, 1976), p. 25.
25. R. E. Frisch, paper presented at the Conference on Nutrition and Human Reproduction, Bethesda, Md., February 1977. [See the conference proceedings, *Nutrition and Human Reproduction*, W. H. Mosley, Ed. (Plenum, New York, 1978), pp. 91–122. See also R. E. Frisch, *Science* **199**, 22 (1978).]
26. J. Bongaarts and H. Delgado, paper presented at the Conference on Natural Fertility held by the International Union for the Scientific Study of Population, Paris, 21–24 March 1977.
27. S. Huffman, K. M. A. Alauddin Chowdhury, W. H. Mosley, "Lactational Amenorrhea: How Is It Affected by Nutritional Status?," unpublished; the findings were presented by Huffman at the annual meeting of the Population Association of America held in St. Louis, 21–23 April 1977.
28. B. Boulier, A. Coale, E. McWilliams. J. Menken, R. Quandt, and J. Tanner suggested several important changes which have improved the clarity of the argument.

6 April 1977; revised 7 October 1977

Trussell's comment (*1*) contains many incorrect statements, both historical and statistical, including the contention of supposed statistical error, as will be shown in detail. Incorrect also is the presumption that over 24 publications on adolescent events and female reproductive ability are to be construed as offering "five pieces of evidence" for the effect of nutrition on fertility. Beginning with the Frisch-Revelle findings on weight at peak velocity and menarche (*2–4*) we followed biological clues and asked biological questions of growth data, using the usual statistical procedures (*2–7*). In Frisch and McArthur (*8*) data on weight and height of women amenorrheic because of weight loss were studied in relation to a minimum weight for height indicated by a fatness index, total water as percent of body weight.

Menarche and fatness: reexamination of the critical body composition hypothesis

J Trussell

Science **200** (4349), 1506-1509.
DOI: 10.1126/science.663636