

Recall of Early Menstrual History and Menarcheal Body Size: After 30 Years, How Well Do Women Remember?

A. Must,^{1,2} S. M. Phillips,¹ E. N. Naumova,¹ M. Blum,³ S. Harris,³ B. Dawson-Hughes,³ and W. M. Rand¹

The validity of recall of early menstrual characteristics is of interest because of their putative role in the etiology of breast cancer and other diseases. A retrospective follow-up of the Newton Girls Study (1965–1975) provided an opportunity to assess the accuracy and precision of recall of several early menstrual characteristics. In 1998–1999, 57 percent of the original 793 Newton Girls Study participants completed a mailed questionnaire to assess the accuracy of recall for age and body size at menarche, usual cycle length during the first 2 years, and age at regularity. Recalled and original age at menarche were highly correlated ($r = 0.79$, $p < 0.001$). The body mass index percentile at menarche was well correlated with recalled body size at menarche ($r = 0.61$, $p < 0.001$), but with some evidence of systematic bias. Overall, a woman's recall of menarcheal age and body size was better than recall of cycle length and occurrence of regularity. The failure to identify certain menstrual characteristics as exposures for subsequent disease may reflect limitations in the accuracy and precision of the recalled measures. *Am J Epidemiol* 2002;155:672–9.

anthropometry; body image; body mass index; menarche; menstruation; recall; somatotypes

Early menstrual and reproductive characteristics are believed to play an important role in a female's subsequent health. These characteristics may serve as exposures or covariates in epidemiologic studies, and their measurement is often based on remote recall, increasing the likelihood of error. To assess the validity of these measures fully, both precision (the absence of misclassification) and accuracy (the absence of bias) should be evaluated when possible.

Menarcheal characteristics of interest in women's health include age at menarche, age at cycle regularity, and menarcheal weight status. Early age at menarche is an established risk factor in the development of breast cancer (1–4) and has also been associated with uterine leiomyomata (5) and endometrial cancer (6).

The number of ovulatory cycles before the first pregnancy has been linked to breast cancer risk (7, 8) and may explain why early menarche increases this risk.

Other menstrual characteristics, such as cycle regularity, appear to be related to the number of ovulatory cycles and, thus, may also represent potential breast cancer risk factors. Extremes of cycle length increase the risk of ovulatory infertility (3). A major source of cycle length variation occurs during the first 2 years after menarche (9–13). Furthermore, girls who establish regular cycles early are more likely to be ovulating than girls who remain irregular.

Physical characteristics, such as weight and height in childhood and adolescence, can also have an impact on adult health. Earlier age at which maximum height is achieved (14, 15) and peak height growth velocity (16) may increase breast cancer risk. Protective effects of elevated preadolescent and adolescent weight status have been reported for breast cancer occurrence (16–18) and for increased adult bone mineral density (19). Obesity present during puberty may be an independent risk factor for other reproductive problems, such as polycystic ovary syndrome and infertility (20). In addition, obesity present in childhood and adolescence represents a major pathway to other adverse health consequences, such as coronary heart disease, type 2 diabetes, and hypertension because of the likelihood that early obesity will persist (21, 22).

Some of the inconsistencies observed in the relation between early menstrual characteristics and long-term consequences may stem from reliance on remote recall. Early menstrual regularity and short or long menstrual cycle length have been associated with breast cancer in some (23–26), but not all (1, 27), studies. Garland et al. (3) suggest that the inconsistent results from case-control studies of the interval between menarche and the establishment of regular cycles in relation to breast cancer may be due to dif-

Received for publication February 26, 2001, and accepted for publication October 2, 2001.

Abbreviations: BMI, body mass index; NGS, Newton Girls Study.

¹ Department of Family Medicine and Community Health, School of Medicine, Tufts University, Boston MA.

² Epidemiology Program, Jean Mayer United States Department of Agriculture Human Nutrition Research Center on Aging, Tufts University, Boston, MA.

³ Calcium and Metabolic Bone Laboratory, Jean Mayer United States Department of Agriculture Human Nutrition Research Center on Aging, Tufts University, Boston, MA.

Correspondence to Dr. Aviva Must, Department of Family Medicine and Community Health, Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111 (e-mail: aviva.must@tufts.edu).

ficuity in recall. Such misclassification would bias effect estimates toward the null. So-called “validation” studies of recall have been undertaken with differing intervals between menarche and recall and with various measures of error (28–32). Much of this research has been done using historical cohorts. Less is known about the validity of other menstrual characteristics, such as cycle length and age at regularity.

The original materials of the Newton Girls Study (NGS) provided the rare opportunity to assess the validity of long-term recall of early menstrual characteristics among a cohort of women nearing menopause. Menstrual cycle data are inherently complex and present the researcher with substantial challenges. To assess the precision and accuracy with which middle-aged women recall their menarcheal age, cycle length, and body size at menarche, we developed and applied a set of analytical approaches.

MATERIALS AND METHODS

Childhood data

This retrospective cohort study uses data from the NGS, a 10-year prospective study of physical growth and sexual maturation initiated in 1965 (33–35). In the original study, of the 1,258 girls in grades 3 and 4 attending public schools in Newton, Massachusetts, who were invited, 793 (63 percent) agreed to participate. Girls were followed through their first menstrual period (menarche) and for approximately 2 years postmenarche. The girls’ mothers were sent monthly questionnaires that requested the report of height, weight, the exact date of each period, the amount of flow, and the presence of discomfort and/or pain for the most recent menstrual period. These contemporaneous childhood reports constitute what we refer to as “original” measurements obtained from NGS data.

Adult data

In 1998, up to 33 years after the original NGS, recontact with the original study subjects was undertaken to collect recalled information on early menstrual characteristics. Efforts to locate NGS participants included the use of the Massachusetts Department of Public Health Vital Statistics Office, alumni books for the Newton high schools, voting books from Newton City Hall, Internet address locators, and credit bureau searches. Because this cohort is composed entirely of women, identifying married names was essential to location efforts. An initial contact letter and response card were sent through the mail to the subject or, if no address for the subject was available, to one of her parents or siblings. Up to three repeat mailings were sent to confirmed addresses. For assurance that the subject was unaware of the hypothesis under study, the letter described the study as an investigation of the relation between early growth and later women’s health outcomes. Once the subject returned a postcard agreeing to participate, she was sent a health questionnaire to be completed and returned. The protocol was approved by the Tufts-New England Medical Center Human Investigation Review Committee.

Of the original 793 participants, 1 percent ($n = 9$) were deceased at the time of follow-up or were unable to respond due to poor health. Of the remaining 784 women, 57 percent ($n = 448$) completed the mailed questionnaire, 6 percent ($n = 44$) declined participation, and the remaining women could not be located ($n = 50$) or did not respond to repeated mailings to a verified address ($n = 242$). The follow-up questionnaire asked women to recall their age at menarche, the age at which their periods became regular, and their body size at menarche by using figure drawings. In addition, the questionnaire asked for recall of usual cycle length during the first 2 years of menstruation. The information provided by these 448 respondents constitutes what we refer to as “recalled” measurements obtained from the follow-up questionnaire.

Variables

A set of variables was developed to allow for the comparison of original and recalled measurements. Original age at menarche was extracted from the original records. Original body mass index (BMI) percentile at menarche was calculated from height and weight reported at menarche in reference to the First National Health and Nutrition Examination Survey (36). Original measures of cycle length were extracted from the original data for subjects with adequate data during the first 2 years. Age at regularity was determined from the original data by applying the definition used by the World Health Organization Task Force on Adolescent Reproductive Health (11). A girl was considered to have reached regularity when the lengths of three consecutive cycles were within a range (maximum – minimum) of 10 days and all three cycles were between 20 and 40 days. The girl’s age at the period marking the first of these three consecutive periods was taken as the age at regularity.

Women were asked to recall their age at menarche and at regularity in years and months. They were also asked to report whether or not their periods ever became regular (at any point in their lives) and, if they became regular, the age at which regularity was reached. Because most women participated in the original study only for a maximum of 20–24 menstrual cycles (approximately 2 years), age at regularity could be assessed only for those women who reached regularity during the first 20–24 cycles postmenarche. In addition, women were asked to recall usual cycle length and regularity during the first 2 years of menstruation. Women were also asked to recall their body size at ages 10, 12, 14, 16, and 18 years by using figure drawings ranked from one to nine (37). The drawing corresponding to the age that was closest to the originally reported age at menarche was taken to reflect the subject’s recalled body size at menarche.

The original and recalled variables described above were used to assess the accuracy and precision with which women recall their early menstrual characteristics. Recalled age at menarche was subtracted from original age at menarche to compute error in recall of menarcheal age. Similarly, recalled average cycle length was subtracted from original mean, median, or modal cycle length to compute error in recall of cycle length. A BMI percentile was established for

each drawing based on expert opinion (38). The percentiles used were developed based on the First National Health and Nutrition Examination Survey, conducted from 1971 to 1973 (39). The subjects in the study were born in 1955–1957 and, thus, would have been early adolescents at the time of the Survey.

To assess the error in recall of body size at menarche, we compared these BMI percentiles associated with the drawing selected based on recall to the subject's BMI percentile calculated from original height and weight recorded at menarche.

Analysis

Data were analyzed using SPSS 9.0 (SPSS, Inc., Chicago, Illinois) and S-Plus (Insightful, Seattle, Washington). Statistical significance was set at $p < 0.05$, and two-sided tests were conducted throughout. Student's t tests were used to compare menstrual characteristics between respondents and nonrespondents.

The analytical techniques selected reflect the nature of the data and our particular goals: to evaluate the potential for random misclassification (lack of precision) and bias (lack of accuracy). Pearson and Spearman correlation coefficients were used to examine the association (likelihood of misclassification) between recalled and original menarcheal age and body size at menarche. We also examined the conditional distribution of original BMI percentile for a given recalled body size by using box plots. Linear regression modeling was used to assess the impact of current BMI and early menstrual characteristics on recall of age and body size at menarche.

The potential for bias was explored by examining the disparity between original and recalled measurements. To assess the difference between original and recalled measures (error), we conducted paired t tests for continuous variables (menarcheal age and age at regularity). To explore how recall error varied across the range of original values, we examined the relation between error in recall of menarcheal age, BMI percentile at menarche, and age at regularity. For each of the measures, we also calculated an absolute error ($|\text{recalled} - \text{original}|$), which reflects its magnitude of disparity and related it to the original value of the measurement. To examine the shape of these relations for error and for absolute error, we used a generalized additive model (40) with a Loess smoother as an additive contributor. We plotted the predicted error (or absolute error) and its 95 percent

confidence interval against the original BMI measurement. For assessment of the measure of central tendency most accurately related to recalled cycle length, error in recall of cycle length (recalled – original) was compared with the original mean, median, and modal cycle length.

RESULTS

Subject characteristics and evaluation of participation bias

The time interval for recall (time between original childhood data collection and completion of the follow-up questionnaire) ranged from 23 to 33 years, with a mean of 29 years (1.4 standard deviation). Original age, height, weight, and BMI percentile at menarche for respondents and nonrespondents are shown in table 1. These childhood characteristics did not differ between respondents and nonrespondents. Because some women did not complete all of the recall questions, it was of interest to assess whether the original measures differed between women who did and those who did not recall individual characteristics. Original measures of menarcheal age, mean BMI percentile at menarche, and mean cycle length did not differ between those who did and those who did not provide recalled information. The data available for the specific analyses are provided in table 2. The percentages for those measures that relate to the first 2 years of menstruation are based on smaller sample sizes due to the requirement of 2 years of original data.

Current anthropometric and sociodemographic characteristics are provided in table 3. At the time of follow-up, the majority of women were married (71 percent), were in their early 40s, were White (97 percent), and had completed college or graduate school (84 percent).

Age at menarche

Original mean menarcheal age (12.93 years, 95 percent CI: 12.81, 13.06) did not differ from recalled mean menarcheal age (12.85 years, 95 percent CI: 12.69, 13.00; paired t test, $p = 0.07$). Linear regression analysis indicated that error in recall was not significantly influenced by age at menarche.

On average, women recalled their menarche as being 0.08 years earlier than their original menarche (95 percent CI: –0.18, 0.01). However, the mean *absolute* error in recall ($|\text{recalled} - \text{original}|$) was 0.62 years on average and varied considerably and nonlinearly across menarcheal ages (figure

TABLE 1. Characteristics of subjects at study entry, 1998 adult follow-up of the Newton Girls Study of 1965–1975

	Responders (completed questionnaire)			Nonresponders (did not complete questionnaire)			<i>p</i> value*
	No.	Mean	(SD)†	No.	Mean	(SD)	
Age at menarche (years)	404	12.9	(1.2)	304	12.8	1.2	NS†
Weight at menarche (kg)	363	46.9	(6.6)	269	47.4	7.5	NS
Height at menarche (m)	364	1.56	(0.06)	269	1.55	0.06	NS
BMI† percentile at menarche	362	49.97	(23.8)	268	54.1	26.1	NS

* t test.

† SD, standard deviation, NS, nonsignificant; BMI, body mass index.

TABLE 2. Number of adult respondents to the questionnaire ($n = 448$) with data on early menstrual and menarcheal weight status characteristics, 1998 adult follow-up of the Newton Girls Study of 1965–1975

	Responders who answered this questionnaire		Subjects with recall data who have original data	
	No.	%	No.	%
Age at menarche (years)	407/448	91	368/407	90
Body size at menarche	379/448	85	339/379	89
Occurrence of regularity ^{*,†}	377/448	84	154/377	41
Recall of early cycle length [†]	239/448	53	159/239	67

* Yes/no, based on World Health Organization definition (see Materials and Methods).

† Calculated only for subjects with 2 years of data postmenarche (see Materials and Methods).

1). The smoothed curve for the absolute error and its 95 percent confidence interval (obtained from a generalized additive model) ranged from about 0.4 to 0.8 years and increased sharply after the average age at menarche (about 12.9 years). The error continues to exceed the average of 0.62 years until about age 16 years.

The precision with which menarcheal age was recalled was explored. Fifty-five percent of the women recalled their age at menarche to within half a year of original menarche; 79 percent were accurate to within 1 year of original menar-

che. Original and recalled menarcheal age were highly correlated (Pearson $r = 0.79$; $p < 0.001$).

Recall of body size at menarche

Figure 2 presents a plot of recalled body size and the corresponding distribution of measured BMI percentile at menarche. Although the possible body sizes ranged from one to nine, no women recalled a body size of greater than seven. More than half of the women who recalled their body size as three were above the 50th percentile BMI, and more than three quarters of those who recalled their body size as four were above the 50th percentile. Variability in BMI percentile at menarche decreased markedly across increasing recalled body sizes. Overall, BMI percentile at menarche was well correlated with recalled body size at menarche (Spearman's $r = 0.61$; $p < 0.001$).

An examination of the relation between error in recall of BMI percentile and original BMI (figure 3) reveals substantial systematic bias. The error ranges from an overestimation of 10 percentile units for the thinnest girls to an underestimation of 30 percentile units for normal and heavier girls. Results from regression analysis showed that adult BMI had no additional influence on recall error.

Occurrence of regularity

Sixty-three percent ($n = 97$ of 154) of subjects with original and recalled data remembered their periods as having been regular, based on the World Health Organization definition (see Materials and Methods). About 96 percent ($n =$

TABLE 3. Characteristics of Newton Girls Study subjects at 1998 adult follow-up

	No.	%	Mean	SD*
Age (years)	448		41.97	(0.76)
Weight (kg)	437		63.5	(13.7)
Height (m)	442		1.63	(0.06)
BMI* (kg/m ²)	436		23.78	(4.89)
Education				
Graduate school	207	46.2		
College	169	37.7		
Some college	48	10.7		
High school	23	5.1		
Less than high school	1	0.2		
Race/ethnicity				
White	434	97.1		
Asian/Pacific Islander	5	1.1		
African American	2	0.4		
Hispanic	1	0.2		
Other	5	1.1		
Marital Status				
Married	318	71.1		
Single/never married	62	13.9		
Separated/divorced	36	8.1		
Significant relationship	24	5.4		
Widowed	7	1.6		

* SD, standard deviation; BMI, body mass index.

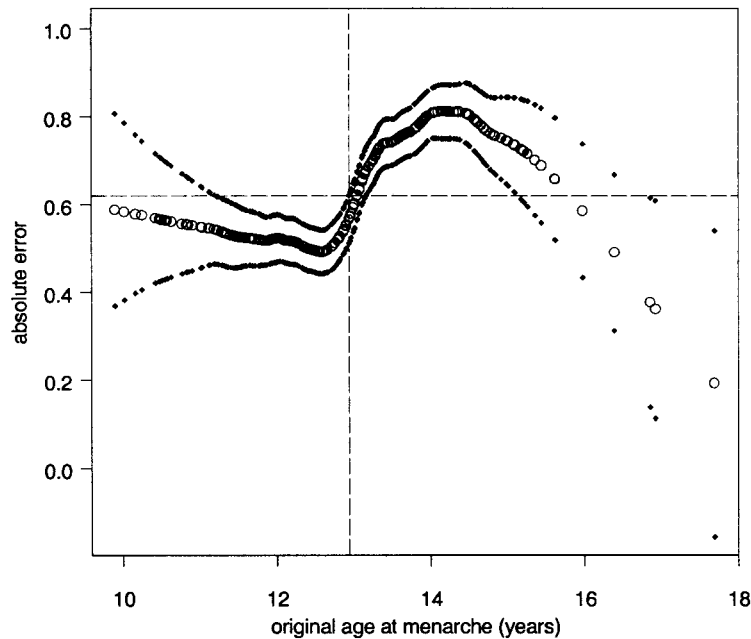


FIGURE 1. The nonlinear relation between original age at menarche and absolute error ($|\text{recalled} - \text{original}|$). Estimates and 95 percent confidence intervals are based on the results of generalized additive modeling, 1998 adult follow-up of the Newton Girls Study of 1965–1975. The horizontal and vertical dashed lines indicate mean absolute error and mean age at menarche, respectively. The middle curve, composed of open circles, indicates estimated absolute error. The upper and lower curves, composed of solid diamonds, indicate the upper and lower confidence limits, respectively.

93 of 97) of women who recalled that their periods were regular met the applied definition of regularity. However, of those who recalled that their early periods were irregular, only 17 percent ($n = 10$ of 57) were actually irregular by the applied definition.

Recall of early cycle length during the first 2 years

Almost all subjects (91 percent) became regular during the first 2 years after menarche. The error in recall of early cycle length was examined for the three basic measures of central tendency: the mean, median, and mode of original cycle lengths over the first 2 years of menstruation. The recall error for each is presented stratified by original and recalled regularity, yielding four estimates (table 4). The three central tendency measures gave very different estimates of recall error. The recall of regularity does not correspond well with the mean or median cycle lengths during the first 2 years as calculated from the original measurements, with recall errors of -9.27 days (mean) and -3.67 days (median) for subjects who recalled being regular. In contrast, for subjects who recalled that their periods were regular during the first 2 years, the recall of cycle length is accurately recalled, with an average error of -0.24 days, when compared with the original modal cycle length.

DISCUSSION

Our efforts to assess the validity of recall of early menstrual characteristics in a cohort of women nearing

menopause were largely successful. Validation studies of the measures of menstrual characteristics are rare, probably because, in addition to being almost prohibitively expensive to acquire, prospectively collected menstrual data are very difficult to work with because of their longitudinal nature and the inevitably large amount of missing data. In this study, the nature of the recalled information did not allow us to evaluate both accuracy and precision for every characteristic of interest. We found that recall was generally better, in terms of accuracy or precision or both, for age at menarche, weight status at menarche, and occurrence of regularity than it was for cycle length. The characteristics for which recall was better were those that relied on memories of specific events and did not require the subject to summarize her experience over a period of time (i.e., questions that asked the subject to summarize her experience over the first 2 years of menstruation).

Our findings with regard to the accuracy of recall of menstrual characteristics confirm and extend the findings reported in other studies (28–32, 41, 42). Previous studies of the recall of age at menarche report correlation coefficients of 0.60–0.81 (28–31, 41), a surprisingly narrow range given that the studies reflect recall intervals of 1–39 years. In those studies for which a mean recall error is reported, mean errors were lower than 0.5 years (28, 30–32, 42), and these small differences were often not different statistically. In our study, we found that recall of age at menarche was generally quite good, both in precision ($r = 0.79$) and accuracy (mean error (recalled – original) = -0.08 years). The average absolute error better describes the degree to which individual measures

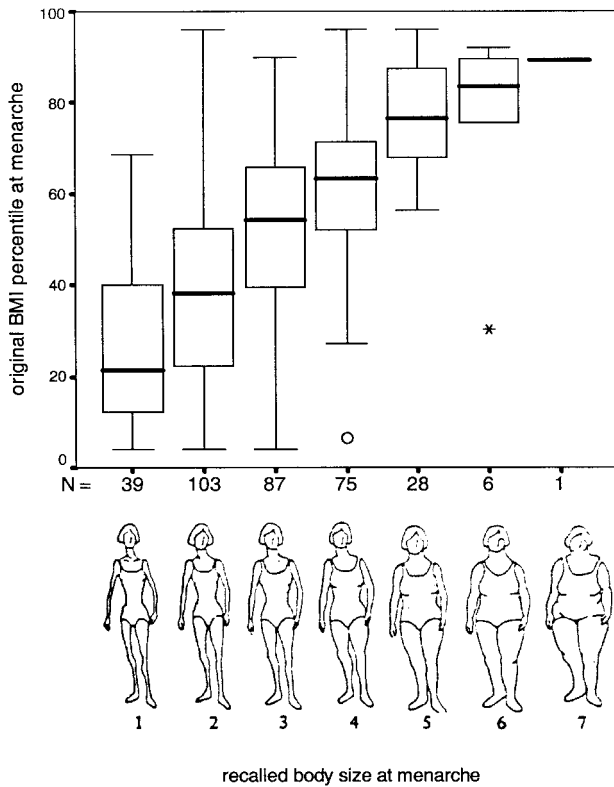


FIGURE 2. The relation of original body mass index (BMI) percentile at menarche versus recalled figure drawing, 1998 follow-up of the Newton Girls Study of 1965–1975. Recalled figure drawings derive from drawings selected to represent body size at ages 10, 12, 14, 16, and 18 years. The drawing that was closest to the original age at menarche was taken to reflect the subject's recalled body size at menarche. Expert median percentiles were 3, 15, 30, 50, 85, 95, and 98, for drawings one to seven, respectively.

deviate from the true measure and will often be more informative than the correlation or mean difference between the recalled and original measures. Further, our study evaluated this error across menarcheal ages and found it to be nonlinear: Absolute error was smallest for menarche that occurred earlier than the mean for the cohort (12.9 years). Error also appeared to decline for older ages at menarche (>16 years), but the sparseness of data at the older ages makes it difficult

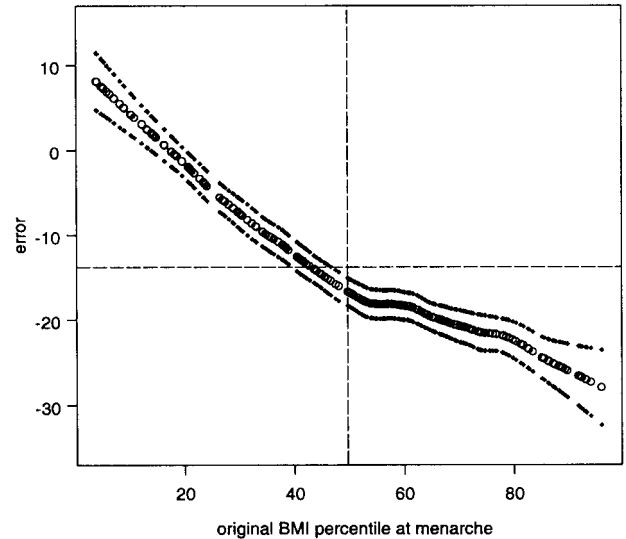


FIGURE 3. The inverse relation between original body mass index (BMI) percentile and error in recall (recalled – original) in which recalled BMI values are based on expert estimates of BMI for a given figure drawing, 1998 adult follow-up of the Newton Girls Study of 1965–1975. Estimates (and 95 percent confidence intervals) are based on the results of generalized additive modeling. Horizontal and vertical dashed lines indicate mean error and mean BMI percentile at menarche, respectively. The middle curve, composed of open circles, indicates estimated error. The upper and lower curves, composed of solid diamonds, indicate the upper and lower confidence limits, respectively.

to draw any conclusion for this group. Given that early menarche is often of interest as a risk factor in studies of cancers of reproductive organs, our data suggest that recalled measures of age at menarche should be quite valid.

Consistent with expectations, research on memory formation indicates that recall ability is directly related to the salience of the event for the respondent (43) and that important personal experiences can produce clear memories of a person's circumstances at the time of the event, so called "flashbulb memories" (44). Our finding that early menarche was well recalled is in keeping with the idea that events that disrupt personal plans aid memory formation. College women who recalled that they were less knowledgeable about menstruation provided more information about the

TABLE 4. Error in recall (recalled minus original mean, median, or mode) of usual cycle length during first 2 years, stratified by original and recalled regularity, 1998 adult follow-up of the Newton Girls Study of 1965–1975

	Recall of regularity	Regular within 2 years			Non regular within 2 years			All subjects		
		No.	Mean error	SD*	No.	Mean error	SD	No.	Mean error	SD
Mean	Yes	65	-8.65	(6.15)	2	-29.27	(27.0)	67	-9.27	(7.76)
	No	30	0.938	(19.8)	7	-4.89	(122.0)	7	-0.166	(52.9)
Median	Yes	65	-3.39	(4.5)	2	-12.7	(9.5)	67	-3.67	(4.84)
	No	30	9.49	(20.3)	7	39.9	(93.7)	7	15.24	(44.0)
Mode	Yes	65	-0.35	(5.4)	2	3.5	(2.1)	67	-0.235	(5.35)
	No	30	14.23	(20.4)	7	74.5	(75.1)	7	25.63	(42.9)

* SD, standard deviation.

first episode, presumably because the event was somewhat traumatic (45).

Recall of body size around the time of menarche is particularly challenging since height and weight are changing during childhood because of growth; subjects are unlikely to recall their height and weight at a given age. Many epidemiologic studies substitute a weight recalled at age 18 or 20 years to reflect adolescent weight status (46, 47). However, if weight status around puberty is the exposure of interest, this compromise may obscure the true relations.

Casey et al. (28) assessed the precision and accuracy of long-term recall of body fatness "relative to peers" during early adolescence for females and found moderately strong correlations ($r = 0.79$) and that the percentage of correct classifications for males and females together (approximately 60 percent) did not vary by maturational timing. Both men and women tended to underestimate their relative fatness during early adolescence. The precision of recalled body size in our study was also fairly good, with a correlation of 0.61. The figure drawings we used in our study are currently used to assess past body size in several large, ongoing prospective studies, including the Nurses' Health Study (48), the Health Professionals Follow-Up Study (49), and the Harvard Alumni Studies (R. S. Paffenbarger, Jr., Harvard School of Public Health, personal communication, 2000). Our previous examination (50) of the validity of these drawings to reflect adolescent BMI recalled after many years yielded similar results, with correlations of 0.65 (at age 10 years) and 0.75 (at age 15 years). Although coefficients of 0.60–0.75 are generally considered to indicate that two measures are well correlated (51), in the recall context, the strength of this relation can be easily misinterpreted. If our interest is to assess the quality of recall across the entire range of values, the regression approach is preferable.

Our available data did not allow us to directly assess bias in recall of body size, but use of expert opinion permits us to explore the nature of recall error. We observed systematic bias: For all but the thinnest girls, subjects underestimated their body size, and the magnitude of the error increased with increasing original BMI percentile.

The literature on the recall of other menstrual characteristics is quite sparse, probably due to the paucity of data sets with the appropriate early measurements. In a sample of 160 subjects in three adult age groupings selected from the Menstrual and Reproductive History Study cohort, several menstrual characteristics, including age at menarche, menstrual cycle length, and cycle variability, recalled during a face-to-face interview were compared with historically recorded data. Cycle length and variability were poorly recalled, with percentage agreements of less than 60 percent for both measures at all ages (32). It is difficult to compare these findings with the current ones because the measure of cycle variability was not comparable with the World Health Organization definition of regularity we used. Furthermore, because the Menstrual and Reproductive History enrolled women during their college years, their earliest recorded periods are for cycles before age 20 years, not for the first 2 years of menstruation, which are generally more variable.

Our evaluation of the concordance of recalled cycle length to original mean, median, and modal cycle lengths suggests that when asked to recall "average cycle length," women report the mode. We conclude that researchers who wish to obtain a measure of cycle length would do well to ask a pair of questions, in which past cycle length is asked only for women who recall that they were regular. The information supplied in this context is apt to reflect the modal cycle length.

This study has a number of noteworthy limitations. First, participation in the original study may have influenced recall. In the original study, all questionnaires and communications were addressed to the mothers to ensure a more reliable and systematic return of the questionnaires and to reduce misreporting due to a girl's desire to be like her classmates. However, later in the study, the girls themselves often provided the information. If participation enhanced recall, one would expect that our findings represent the best-case scenario for remotely recalled measures of early menstrual characteristics. We did exclude data for one study subject who provided us with a lifetime log of every menstrual period she had experienced since menarche. Unfortunately, we did not have data on comorbidity during menarche, which could have influenced the timing of onset or patterns. In addition, the subjects in this study were White, well-educated women, so our findings may not be directly generalizable to women of other races or with lower educational attainment. However, we believe that this inquiry has particular relevance to women's health issues because it is this age group of women who are entering a life stage of increased risks for breast cancer, osteoporosis, and cardiovascular diseases. Furthermore, our study population was predominantly Jewish (based on maiden surname) and professional, both of which are characteristics associated with higher breast cancer risk (52).

In conclusion, epidemiologists interested in menstrual data as exposure measures will probably continue to rely on retrospectively collected data, given the expense and subject burden associated with gathering such data prospectively. Our experience with this rich data set suggests that long-term recall of age and relative weight status at menarche are quite valid as exposure measures. We believe that information on cycle length can be validly obtained only for women who report that they are regular. Our analyses sought to emphasize the importance of assessing both misclassification and bias and their independent contributions to validity.

ACKNOWLEDGMENTS

Supported by the Massachusetts Department of Public Health Breast Cancer Research Program and the Epidemiology Core of the Boston Obesity Nutrition Research Center (grant DK46200).

REFERENCES

1. Titus-Ernstoff L, Longnecker MP, Newcomb PA, et al. Menstrual factors in relation to breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 1998;7:783–9.

2. Sellers TA, Kushi LH, Potter JD, et al. Effect of family history, body-fat distribution, and reproductive factors on the risk of postmenopausal breast cancer. *N Engl J Med* 1992; 326:1323-9.
3. Garland M, Hunter DJ, Colditz GA, et al. Menstrual cycle characteristics and history of ovulatory infertility in relation to breast cancer risk in a large cohort of US women. *Am J Epidemiol* 1998;147:636-43.
4. Kelsey JL, Gammon MD, John EM. Reproductive factors and breast cancer. *Epidemiol Rev* 1993;15:36-47.
5. Marshall L, Spiegelman D, Goldman M, et al. A prospective study of reproductive factors and oral contraceptive use in relation to the risk of uterine leiomyomata. *Fertil Steril* 1998; 70:432-9.
6. McPherson CP, Sellers TA, Potter JD, et al. Reproductive factors and risk of endometrial cancer: the Iowa Women's Health Study. *Am J Epidemiol* 1996;143:1195-202.
7. Colditz GA, Frazier LA. Models of breast cancer show that risk is set by events of early life: prevention efforts must shift focus. *Cancer Epidemiol Biomarkers Prev* 1995;4:567-71.
8. Henderson BE, Ross RK, Judd HL, et al. Do regular ovulatory cycles increase breast cancer risk? *Cancer* 1985;56:1206-8.
9. Lee I, Paffenbarger RS Jr. Change in body weight and longevity. *JAMA* 1992;268:2045-9.
10. Treloar AE, Boynton RE, Behn BG, et al. Variation of the human menstrual cycle throughout reproductive life. *Int J Fertil* 1970;12:77-126.
11. World Health Organization. World Health Organization Multicenter Study on Menstrual and Ovulatory Patterns in Adolescent Girls. *J Adolesc Health Care* 1986;7:236-44.
12. Golub S. Periods: from menarche to menopause. Newbury Park CA: Sage Publications, Inc, 1992.
13. Harlow SD, Ephross SA. Epidemiology of menstruation and its relevance to women's health. *Epidemiol Rev* 1995;17:265-86.
14. La Vecchia C, Negri E, Parazzini F, et al. Height and cancer risk in a network of case-control studies from northern Italy. *Int J Cancer* 1990;45:275-9.
15. Li CI, Malone KE, White E, et al. Age when maximum height is reached as a risk factor for breast cancer among young U.S. women. *Epidemiology* 1997;8:559-65.
16. Berkey CS, Frazier AL, Gardner JD, et al. Adolescence and breast carcinoma risk. *Cancer* 1999;85:2400-9.
17. Stoll BA. Teenage obesity in relation to breast cancer risk. *Int J Obes Relat Metab Disord* 1998;22:1035-40.
18. Huang Z, Hankinson SE, Colditz GA, et al. Dual effects of weight and weight gain on breast cancer risk. *JAMA* 1997;278: 1407-11.
19. Blum M, Must A, Harris SS, et al. Association of documented weight at menarche with premenopausal bone mineral density. (Abstract). *J Bone Miner Res* 1999;14:S379.
20. Lake JK, Power C, Cole TJ. Women's reproductive health: the role of body mass index in early and adult life. *Int J Obes* 1997;21:432-8.
21. Must A, Strauss RS. Risks and consequences of childhood and adolescent obesity. *Int J Obes Relat Metab Disord* 1999;23 (suppl 2):S2-11.
22. Serdula MK, Ivery D, Coates RJ, et al. Do obese children become obese adults? A review of the literature. *Prev Med* 1993;22:167-77.
23. Apter D, Vihko R. Early menarche, a risk factor for breast cancer, indicates early onset of ovulatory cycles. *J Clin Endocrinol Metab* 1983;57:82-6.
24. Vihko R, Apter D. Endocrine characteristics of adolescent menstrual cycles: impact of early menarche. *J Steroid Biochem* 1984;20:231-6.
25. Magnusson CM, Persson IR, Baron JA, et al. The role of reproductive factors and use of oral contraceptives in the aetiology of breast cancer in women aged 50 to 74 years. *Int J Cancer* 1999;80:231-6.
26. Whelan EA, Sandler DP, Root JL, et al. Menstrual cycle patterns and risk of breast cancer. *Am J Epidemiol* 1994;140:1081-90.
27. Adami HO, Bergstrom R, Lund E, et al. Absence of an association between reproductive variables and the risk of breast cancer in young women in Sweden and Norway. *Br J Cancer* 1990;62:122-6.
28. Casey VA, Dwyer JT, Coleman KA, et al. Accuracy of recall by middle-aged participants in a longitudinal study of their body size and indexes of maturation earlier in life. *Ann Hum Biol* 1991;18:155-66.
29. Bergsten-Bruceford A. A note on the accuracy of recalled age at menarche. *Ann Hum Biol* 1976;3:71-3.
30. Livson N, McNeill D. The accuracy of recalled age at menarche. *Hum Biol* 1962;34:218-21.
31. Damon A, Bajema CJ. Age at menarche: accuracy of recall after thirty-nine years. *Hum Biol* 1974;46:381-4.
32. Bean JA, Leeper JD, Wallace RB, et al. Variations in the reporting of menstrual histories. *Am J Epidemiol* 1979;109:181-5.
33. Zacharias L, Rand WM. Adolescent growth in height and its relation to menarche in contemporary American girls. *Ann Hum Biol* 1983;10:209-22.
34. Zacharias L, Rand WM. Adolescent growth in weight and its relation to menarche in contemporary American girls. *Ann Hum Biol* 1986;13:369-86.
35. Zacharias L, Rand WM, Wurtman RJ. A prospective study of sexual development and growth in American girls: the statistics of menarche. *Obstet Gynecol Surv* 1976;31:328-37.
36. Must A, Dallal GE, Dietz WH. Reference data for obesity: 85th and 95th percentiles of body mass index (wt/ht²) and triceps skinfold thickness. *Am J Clin Nutr* 1991;53:839-46.
37. Stunkard AJ, Sorensen T, Schulsinger F. Use of the Danish adoption register for the study of obesity and thinness. In: Kety SS, Rowland LP, Sidman SW, et al., eds. The genetics of neurological and psychiatric disorders. New York, NY: Raven Press, 1983:115-20.
38. Must A, Phillips SM, Stunkard AJ, et al. Expert opinion on body mass index percentiles for figure drawings at menarche. *Int J Obes (In press)*.
39. Statistics NCFH. Plan and operation of the Health and Nutrition Examination Survey, United States, 1971-1973. Vital and Health Statistics 1973:1.
40. Hastie TJ, Tibshirani RJ. Generalized additive models. New York, NY: Chapman & Hall, 1990.
41. Damon A, Damon ST, Reed RB, et al. Age at menarche of mothers and daughters with a note on the accuracy of recall. *Hum Biol* 1969;41:161-75.
42. Koo MM, Rohan TE. Accuracy of short-term recall of age at menarche. *Ann Hum Biol* 1997;24:61-4.
43. Jobe JB, Tourangeau R, Smith AF. Contributions of survey research to the understanding of memory. *Appl Cogn Psychol* 1993;7:S67-84.
44. Conway MA. Flashbulb memories. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc, 1995.
45. Pillemer DB, Koff E, Rhinehart ED, et al. Flashbulb memories of menarche and adult menstrual distress. *J Adolesc* 1987;10: 187-99.
46. Ravelli AC, van der Meulen J, Osmond C, et al. Obesity at the age of 50 y in men and women exposed to famine prenatally. *Am J Clin Nutr* 1999;70:811-16.
47. Camargo CA, Weiss ST, Zhang S, et al. Prospective study of body mass index, weight change, and risk of adult-onset asthma in women. *Arch Intern Med* 1999;159:2582-8.
48. Berkey CS, Frazier AL, Gardner JD, et al. Adolescence and breast carcinoma risk. *Cancer* 1999;85:2400-9.
49. Munoz KA, Ballard-Barbash R, Graubard B, et al. Recall of body weight and body size estimation in women enrolled in the Breast Cancer Detection and Demonstration Project (BCDDP). *Int J Obes Relat Metab Disord* 1996;20:854-9.
50. Must A, Willett WC, Dietz WH. Remote recall of childhood height, weight, and body build by elderly subjects. *Am J Epidemiol* 1993;138:56-64.
51. Colton T. Statistics in medicine. New York, NY: Little, Brown & Co, 1974.
52. Kelsey JL. The epidemiology of breast cancer. *CA Cancer J Clin* 1991;41:146-65.