Longitudinal, epidemiologic studies of female reproductive aging

Rebecca J. Ferrell¹ and MaryFran Sowers²

¹National Institute on Aging, National Institutes of Health, Bethesda, Maryland. ²Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan

Address for correspondence: Rebecca J. Ferrell, Scientific Review Branch, National Institute on Aging, National Institutes of Health, 7201 Wisconsin Avenue, Room 2C212, Bethesda, Maryland 20892. rebecca.ferrell@nih.gov

Human female reproductive aging consists of multiple processes and interacts with other physiological systems in unique ways. Here we discuss eight recent longitudinal, epidemiologic studies of female reproductive aging that include endocrine data to highlight their contributions to our understanding of these various aging processes and their interactions. Specifically, we review data on ovarian and nonovarian reproductive aging processes and reproductive staging. We consider these data in the context of longitudinal research design and research goals, identify limitations of the studies but also ways in which existing longitudinal data can further our understanding of aging processes, and make recommendations for future studies of female reproductive aging.

Keywords: ovarian aging; hypothalamic-pituitary axis; menstrual cycle; epidemiologic

Introduction

Understanding the relationship between female reproductive aging and more general aging processes is uniquely challenging in modern humans. Few species and none of our primate relatives have postreproductive intervals as long, relative to total lifespan, as those observed in human females.¹ Reproductive aging in women appears markedly accelerated against the backdrop of the elongated human lifespan and the more gradual aging processes observed in other physiological systems.

Interest in female reproductive aging is not only evolutionarily based but reflects a desire to address health concerns that women face at various stages of reproductive and postreproductive life. Changes in fertility, menopausal symptoms, and health risks during and after the menopausal transition are primary concerns, and research advances continue to inform decisions about appropriate treatments and interventions.² Assessing the interaction among various aging processes is a necessary part of this research effort. Reproductive aging itself can include multiple processes that occur at different rates and whose interactions are not fully understood. Elevated follicle-stimulating hormone (FSH), for example, may be a function of final depletion of ovarian follicles (and associated steroid hormones) at midlife, which in turn releases gonadotropins from negative feedback, but also could be the result of gradual changes in hypothalamic-pituitary sensitivity to this negative feedback.¹³⁻⁵ Also, it is not always clear whether a particular health risk or outcome is the result of reproductive changes, more generalized age-related processes, or both.⁴

Addressing these research challenges requires a broad approach. Epidemiologic studies represent an important component of the research effort because they inform us about female reproductive aging in population-based samples. The first large-scale studies of female reproductive function, in the mid-20th century, were epidemiologic in design and focused on describing the “normal” menstrual cycle changes with age.⁶⁻⁷ With the development of modern assay methodologies that could be implemented in populations, researchers established normative hormone profiles in the menstrual cycle⁸⁻⁹ as part of a larger effort to understand the proximate reproductive functions of the hypothalamic-pituitary-ovarian axis. Concurrently, clinical researchers were...
identifying conditions in which “normal” reproductive function was in some way compromised, in an effort to understand and treat infertility. And observations of postmenopausal estrogen loss, associated health risks (e.g., osteoporosis, cardiovascular risk), and the potential for hormone replacement therapy to reduce those risks, became a primary focus area. These research efforts laid the groundwork for more recent longitudinal, epidemiologic studies to characterize female reproductive aging in community samples and to reflect the sources of variation that may be contributed by race/ethnicity, body size, or behavioral practices. The studies vary in terms of research goals and therefore in hormonal, menstrual, and covariate measures, but the projects together provide an important body of data. In this chapter, we focus on eight of these studies, specifically those that have assessed reproductive hormone changes with age. Several of these studies have been reviewed elsewhere as parts of broader discussions of female reproductive aging, including the relationship and interactions between reproductive aging and more generalized aging processes. Here, we consider the findings of these studies in terms of ovarian and nonovarian aging processes, research goals and epidemiologic study design, and suggestions for future research efforts.

**Study summaries**

The following is a summary of longitudinal, epidemiologic studies of reproductive aging that have been established within the last 25 years, and have endocrine measures, reasonable retention, and sufficient duration of follow-up to describe important physiological processes. In most, but not all cases, the age range of participants reflects a focus on perimenopause. Many of the studies are based on cross-sectional studies in order to establish some degree of generalization to community samples.

There is a notable contrast across studies in hormone sampling, which reflects differences in overall research goals. Those studies that aim to develop a broad, global understanding of the menopausal transition and collect data on a wide range of variables, including social, behavioral, and physiological factors, tend to have annual or monthly hormone measures that can be analyzed in conjunction with many other variables over a number of years. Annual hormone measures are taken in the early to middle follicular phase, which is easily identifiable relative to the menstrual period. Such measures can not, however, provide information about ovulatory status or within-cycle hormone patterns, and may or may not be comparable to other studies that sampled on slightly different days of the follicular phase. Those studies that focus more on the physiological, mechanistic underpinnings of reproductive aging generally take more frequent (e.g., weekly or daily) samples which allow for analyses of changes in within- and across-cycle hormone dynamics (e.g., peak progesterone, luteinizing hormone [LH], or FSH occurring outside the follicular phase) and ovulatory status. The drawbacks of an intensive collection routine are financial and time constraints, participant burden, and (usually) lower participant numbers compared with annual-measure studies.

With these factors in mind, the studies are described here in the time order of their development, and key features are compared in Figure 1.

Massachusetts Women’s Health Study (MWHS)—longitudinal component: 1986–1991

The goal of the MWHS was to describe how women respond to the menopausal transition, and to determine how social, behavioral, and physiological factors affect women’s experiences of the transition. Initially, white women (n = 8050) from the greater Boston area who were not postmenopausal (<11 months amenorrhea) were followed from 1981 to 1986 (ages 45–55 at baseline). Multiple questionnaires covering menstrual, health, lifestyle, and sociodemographic characteristics were administered. A second component of the study (MWHS II; 1986–1991; n = 427 from original cohort) included annual collection of biomarker data. These participants (ages 50–60 years) had an intact uterus, at least one ovary, and no more than 11 months of amenorrhea. Annual blood samples (from day 5 to 7 of menstrual cycle) were assayed for estradiol (E2) and FSH. Data were excluded if women used exogenous hormones or were missing predictor or outcome measures.

Specific topics addressed in the MWHS included menopausal symptoms, sexual function, depression, healthcare utilization, weight, cardiovascular health, women’s attitudes about menopause, and reproductive hormone patterns. The study was particularly important in demonstrating for...
the first time and in a community-based sample the generally positive attitudes of women in the menopausal transition, and that some negative experiences at midlife may not be menopause-specific but rather age-related psychosocial and physiological changes. The annual biomarkers allowed for additional longitudinal analyses of associations among biological reproductive hormones and other biological and psychosocial factors, such as depression, weight, and menstrual cycle characteristics.

Melbourne Women’s Midlife Health Project (MWMHP)—longitudinal component: 1991–2004

The MWMHP was similar to the MWHS in addressing a variety of physiological, social, and behavioral factors associated with the menopausal transition. This study also focused more specifically on hormonal changes of the transition.

A cross-sectional sample of 2001 Australian-born white women (45–55 years of age) was the sampling frame for the study’s longitudinal component.
The longitudinal study included 438 women who initially had menstruated within 3 months of examination and were without exogenous hormone use. Women participated in annual health, menstrual, and hormone data collection over a subsequent 13-year period. Annual blood samples were collected (from days 4 to 8 of menstrual cycle or after 3 months of amenorrhea) and assayed for E2, FSH, inhibin A, inhibin B, immunoreactive total inhibin, testosterone (T), and sex hormone-binding globulin (SHBG). Factors addressed by the MWMHP included mood, sexual function, cholesterol levels, bone density, weight, hormone therapy, well-being and quality of life, menopausal symptoms, cardiovascular health, diet, and cognition. The MWMHP was one of the first studies to show long-term changes in a number of reproductive hormones and in menstrual cycle characteristics, relative to the timing of the final menstrual period (FMP). It also identified associations between reproductive hormones and variables, such as menopausal symptoms, menstrual cycle characteristics, sexual function, and lipids and blood pressure.

Seattle Midlife Women’s Health Study (SMWHS)—longitudinal component: 1996–2005

The SMWHS, initiated in 1990, also addressed a wide range of menstrual and health factors potentially associated with the menopausal transition. A multi-ethnic community sample (recruited by telephone) of women (n = 367, ages 35–55 years, median age 41 at baseline, 20% nonwhite [black, Asian, Native American]) was characterized with annual health and more intensive menstrual data by diary. Commencing in 1996, a subsample of women (n = 170) was selected for a more detailed evaluation of hormonal changes. Women were excluded if they had not menstruated in the past 12 months, had both ovaries removed, or were unable to speak or understand English. Eligible women provided a single “day 6” urine specimen for eight to 12 menstrual cycles per year from 1996 to 2000 and quarterly from 2001 to 2005. Urine samples were assayed for estrone glucuronide (E1g), follicle-stimulating hormone (uFSH), testosterone (uT), and cortisol (uC).

Specific topics addressed by the SMWHS included women’s perceptions and expectations of menopause, mood and depression, memory, menopausal symptoms, hormone therapy, well-being, stress, and menopausal stages, for example, Refs. 22 and 23. Endocrine data have been assessed in relation to factors, such as menopausal symptoms and stages, and genetic polymorphisms in the estrogen pathway. The wider age range and multi-ethnic nature of the participants in this study was also different from the earlier studies.

Michigan Bone Health and Metabolism Study (MBHMS)—longitudinal study: 1992–present

The MBHMS is a population-based longitudinal study with a primary focus on the relationships among reproductive hormones, musculoskeletal and metabolic diseases, and health risks. A total of 664 white women (24–44 years at baseline) were recruited using two sampling frames to generate the population census from which 81% were successfully recruited. Participants have provided annual blood and urine specimens (days 2–7 of menstrual cycle, or on the anniversary of study enrollment for those participants with increasingly irregular cycles) and menstrual and health data. Among a range of other analytes, serum has been assayed for E2, T, SHBG, FSH, anti-Müllerian hormone (AMH), and inhibin B.

The MBHMS has investigated potential relationships between reproductive hormones and factors, such as bone mineral density, osteoarthritis, physical activity, body composition, and metabolic biomarkers, for example, Refs. 27–29, and has also modeled changes in estrogen and FSH relative to the timing of menopause. The study is also notable in having the youngest age range at baseline of all the studies discussed here, allowing it to more fully address the transition though reproductive aging into the postmenopause.

Study of Women’s Health Across the Nation (SWAN)—longitudinal component: 1996–present

SWAN is investigating social and biological factors associated with aging and the menopausal transition. It is one of the most comprehensive of the studies described here, in the sense that it attempts to measure a large number of variables and potential confounding factors, and to collect both annual and daily endocrine data. The initial longitudinal study was based on a cross-sectional survey of a multi-ethnic, community-based sample of women.
From this survey, a subset of 3-302 women (ages 42–52 at baseline) was selected. To be enrolled in the study, women were still menstruating, had an intact uterus and at least one ovary, and could not be pregnant or taking exogenous hormones. Enrollees self-identified as African-American (25%), Caucasian (50%), Chinese (8%), Hispanic (8%), and Japanese (9%). There have been assays of early follicular phase FSH, E2, T, SHBG, and dehydroepiandrosterone sulfate (DHEAS) in the annual serum collections. Interview and diary data included information about each woman’s menstrual cycle.

A multi-ethnic subset of 848 women are taking part in the longitudinal Daily Hormone Study (DHS) in which women collect daily first-morning urine samples during one menstrual cycle (or an interval of 50 days), annually; these urine specimens are analyzed for estrone conjugates (E1c), pregnanediol glucuronide (PdG), uFSH, and uLH.

The topics investigated in SWAN are extensive and include analyses of health, psychosocial, and biomarker data. The endocrine data for SWAN are unique because they include, for an unprecedented number of women in a multi-ethnic sample, both annual and daily (for one menstrual cycle per year) data. The annual data, for example, have contributed to our understanding of long-term changes in hormones with reproductive age. The daily data allow for analyses of potential associations among health and psychosocial factors (e.g., hypothalamic-pituitary function, sleep, bone mineral density, cardiovascular risk) and within-cycle hormone patterns and ovulatory status, for example, Refs. 30, 32, 33, and 35–40.

The Pennsylvania Ovarian Aging Study (POAS)—longitudinal component: 1996–present

The POAS was designed to “identify hormonal, clinical, behavioral, and demographic factors associated with ovarian aging” (p. 544). Participants were 436 randomly selected black and white women from Philadelphia County (35–47 years at baseline) who had had a regular menstrual cycle (22–35 days long) within the 3 months prior to enrollment. Single day hormone data (day 1–6 of menstrual cycle) were collected four times between 1996 and 2004, with approximately 2 years between collection times. Women also kept a record of menstrual cycles. Women had an intact uterus and at least one ovary and were absent health states that might affect ovarian or hormonal function. Blood samples were assayed for E2, FSH, LH, inhibin B, and DHEAS.

The POAS focuses on the relationship between endocrine changes related to ovarian aging (see further) and a variety of potentially associated factors, including menopausal symptoms, sleep, depression, weight gain, physical activity, menstrual bleeding patterns, and sexual function, for example, Refs. 43–46.

The Biodemographic Models of Reproductive Aging (BIMORA) project—longitudinal study: 1998–2002

The BIMORA project was developed to assess the hormonal and menstrual changes occurring with reproductive aging and to understand how these changes are related to the process of follicular depletion, whereby the ovary is ultimately depleted of ova and their associated steroid hormones. Enrollees were 156 white women (ages 25–58; mean 47.6 years at baseline) drawn from participants in the TREMIN Research Program on Women’s Health. Enrollees collected daily urine samples for 6 months (January–July) for 5 consecutive years (1998–2002) while continuing to provide menstrual diary data. At baseline, women had to have at least one ovary, and could not be pregnant, breastfeeding, or using exogenous hormones. The urine specimens were assayed for E1g, PdG, uFSH, and uLH.

Similar to SMWHS, MBHMS, and POAS, the participant age range of the BIMORA study was wide, with a low minimum age at baseline, to capture any endocrine changes that might occur earlier in adult reproductive life. The BIMORA project was the first to collect daily data for multiple consecutive menstrual cycles from each participant. This type of data allowed for the modeling of individual 6-month trajectories of steroid and gonadotropin hormones over as many as 5 consecutive years. It has also allowed for analyses of within-cycle hormone dynamics, including frequency of anovulation with age, and changes in progesterone, estrogen, and unopposed estrogen with age or menopausal stage.

The FREEDOM study is similar to the BIMORA study in that it was designed to understand changes in reproductive hormones and menstrual cycle dynamics at the daily level. The sample was 112 white British women (ages 30–58; median 44 years at baseline) volunteers. They collected daily urine samples and menstrual bleeding data for 6–18 months, consecutively. Urine specimens were assayed for E1g, PdG, uFSH, and uLH. Women were excluded if they were pregnant or breastfeeding, using exogenous hormones, or had pituitary disorders. The FREEDOM study has demonstrated important relationships among the reproductive hormones and cycle characteristics within and across consecutive menstrual cycles, including the relationship between FSH, follicular phase length, and age and the lengthening of the follicular phase associated with estrogen that remains low for a portion of the follicular phase.53

Hormonal changes with reproductive age

Epidemiologic studies provide an opportunity to evaluate hormonal changes at both the individual and population levels, and the eight studies just described have variously extended our understanding of some of the key hormonal changes that occur during female reproductive aging.

Declining inhibin B and AMH (products of the ovarian follicles) are emerging as some of the earliest indicators of approaching perimenopause. These declines, which occur during late premenopause when menstrual cycles are still regular, are thought to reflect the dwindling pool of ovarian follicles. Furthermore, concurrently increasing FSH may reflect a decrease in negative feedback control by these ovarian hormones. The MWMHP was one of the first studies to describe this late premenopausal decline in inhibin B alongside a slight rise in FSH, in a population sample. Findings from the POAS and the MBHMS were similar and the latter showed that AMH declines to values below detection 5 years before the FMP.

When variation in menstrual cycle length starts to occur during early perimenopause, a combination of elevated FSH and maintained or elevated estrogen has been observed, which may indicate that remaining ovarian follicles are being hyperstimulated by elevated FSH. In the MWHS, cohort E2 averages did not change significantly across three annual follow-up measures, even though the percentage of premenopausal women decreased through time. The SMWHS and the FREEDOM study found that E1g remained stable during early perimenopause. The MWMHP found that FSH increased up to and beyond the FMP and then leveled off, while E2 was sustained until 2 years before the FMP. In the MBHMS, population-average E2 remained stable from −10 to −2 years before the FMP. Elevated estrogen has been observed in the BIMORA study and in earlier work by some of the SWAN investigators in which daily hormone data could be analyzed for complete menstrual cycles. Variability in estrogen within and across cycles, the age ranges of participants, and differences in sampling schemes may partly explain the different results across studies.

Rapid declines in estrogen occur during late perimenopause (when variable menstrual cycle length becomes more pronounced), largely in the 2 years before the FMP. This decline is thought to reflect the imminent depletion of estrogen-producing ovarian follicles. Several of the epidemiologic studies have identified this estrogen decline at the population or individual levels. Annual measures have the advantage of being able to characterize the drop in estrogen relative to the FMP, and can estimate the rates of the decline. Daily measures are informative in showing trajectories of declining estrogen for individual women.

Some menopausal changes may extend beyond the process of follicular depletion. Increasing FSH may be in part the result of an aging hypothalamic-pituitary axis in which gonadotropin pulsatility changes and the responsiveness to ovarian steroids decreases. Also, indirect evidence suggests the possibility of reduced LH sensitivity to estrogen: although cycle-average estrogen levels are lower for anovulatory cycles, some anovulatory cycles have follicular phase estrogen levels similar to ovulatory cycles yet do not elicit an LH surge.

Other age-related changes to the oocytes and ovarian tissues may also be affecting cell function, including chromosomal mutations, decreasing telomere length, and accumulating metabolic debris. It remains understudied as to how environmental factors affect declining oocyte quality or ovarian tissues, or how these changes interact with...
the process of follicular depletion. While direct measurements of such changes at the cellular and tissue levels are challenging in longitudinal epidemiologic studies, these studies may generate indirect evidence for relationships among oocyte quality, ovarian tissue function, and follicular depletion. For example, smoking is known to affect egg and ovary quality\(^6\) and has also been shown (Sowers et al.\(^{63}\)) to be associated with earlier menopause and faster rates of AMH decline.

**Defining reproductive stages**

At late reproductive ages, generally menstrual cycles become shorter, and then cycle length variability increases during perimenopause.\(^6\) On the basis of such changes, many of the epidemiologic studies have used series of menstrual bleed data (in some cases complemented by hormone values) to establish population-based reproductive staging paradigms.\(^{20,25,26,30,54,64–72}\) Certain levels of change from the menstrual cycle “norm” of earlier reproductive ages are defined as marking entry into the next stage of reproductive aging. In this way, data can be grouped for women of similar biological age regardless of their chronological ages. The high level of variation in aging trajectories and age at menopause also makes staging systems particularly valuable for giving women an idea of their proximity to menopause. It should be noted, however, that the common staging systems in use today have been developed in large part from consensus discussions, not all women go through all stages or expected sequences of stages, and FMP cannot be precisely predicted by any staging system.

The Stages of Reproductive Aging Workshop (STRAW) reproductive staging system\(^{66}\) consists of seven stages. The mid- to late reproductive stages are indicated by regular cycles and lower FSH, followed by regular cycles and elevated FSH. In the perimenopausal stages, FSH remains elevated and cycles change from variable length to intermenstrual intervals greater than 60 days. Recent analyses\(^{69,70}\) have applied the STRAW system to multiple longitudinal menstrual cycle diary data (MWMHP, TREMIN, SMWHS, SWAN) in order to test the utility of the stages for determining entry into early and late perimenopause, and for determining time to FMP. The use of cycle variability >6 days for defining early perimenopause, and the use of a first >60 day intermenstrual interval for entering late perimenopause, were supported, but additional research is needed to identify more precise markers of the early perimenopause.

Multiple hormone values or hormonal patterns across a menstrual cycle have also been proposed for refining reproductive stages, since it is clear that single hormone values (e.g., FSH) do not provide significantly different information about proximity to the FMP than menstrual cycle changes.\(^{65,69}\) Miro et al.\(^{54}\) proposed a five-stage classification based on FSH and menstrual cycle characteristics that differs from STRAW by defining an FSH threshold for each stage and by including in the definitions cycles with delayed follicular development and anovulation. This method may be difficult to apply in clinical settings where, for example, anovulatory cycles can not easily be identified. Longitudinal analyses of annual FSH\(^{25}\) and E2\(^{26}\) measures in the MBHMS study have shown systematic rates of change and acceleration of change in FSH and E2 that can be characterized in definable stages in the years around the FMP. It should be noted that even these multisample methods do not provide precise information about time to FMP.

**Limitations and future research**

Researchers are cognizant of several issues that limit the analyses that can be done with longitudinal data. First, age ranges of participants in most of the studies described were in the 40s and early 50s, and consideration of those early or mid-late reproductive events that entrain events of the late reproductive age were thereby excluded from study. Second, epidemiologic studies may be constrained to incorporate exclusion criteria that preclude certain population segments (e.g., ethnicity, hormone therapy users, early or late entrants into a particular reproductive stage), thus limiting generalizability of results. Finally, as discussed above, the hormone sampling strategies used in longitudinal studies vary, from large samples of annual measures to smaller samples of daily hormone measures, thus limiting the types of analyses that can be done in each type of study.

One of the goals in female reproductive aging research continues to be to understand attributes leading up to and following the FMP. For studies with annual hormone data, additional data can be pursued using archival specimens, including genetic data, for additional assays thought to be associated with
reproductive aging. It is possible that a combination of hormonal and genetic markers from a single annual specimen, along with specific covariate and menstrual cycle data, could improve upon current reproductive staging efforts. To achieve this, however, specimens and suitable analytical techniques must be available.

For studies with daily hormone data, existing data on full and sequential menstrual cycles could be used to develop sampling strategies for predicting time to FMP and to identify subtle relationships among hormone changes across the various parts of the menstrual cycle. Indeed, a main advantage of existing daily hormone data is that they allow for analyses of sampling strategies or hormonal relationships outside of the early follicular phase. Because daily data span the menstrual cycle, however, external markers, apart from actual menstrual bleeding, would be required to relate to the timing of the various parts of their cycle (e.g., initial estrogen rise, ovulation, luteal phase). The daily hormone studies could also add genetic and additional hormone assay data from archival specimens, but these may require additional analytical work to understand how the urinary excretion levels are related to cellular and molecular attributes.

The combination of rich longitudinal data from the perimenopause and postmenopause could be invaluable for showing associations between perimenopause, health outcomes (e.g., osteoporosis, cardiovascular disease), and other aging processes. The health-related outcomes will need to be common because the sample sizes of most studies of longitudinal aging preclude having sufficient health events to identify, for example, cancers, autoimmune disorders for which women are disproportionately affected, or neurological impairments that may disproportionately affect women.

Genetic markers are likely to become an increasingly common tool in longitudinal studies. Already, some longitudinal studies have examined relationships between specific polymorphisms in the estrogen pathway and later reproductive and health outcomes. Markers for differences in follicular depletion rate and markers for specific disease states are likely to grow in number, and it will be important for existing studies with specimens to consider if and how additional genetic markers can inform their studies. More generally, the use of repositories for specimens and data sets will become increasingly important. The longitudinal studies are in various stages in terms of making their data available to others; most of those studies for which data are not yet available to others are still in the active analysis phase. Others may require additional resources to engage in broad data sharing.

Existing longitudinal studies can play an important role in informing future study designs in human female reproductive aging, through an appreciation of both their limitations and their accomplishments. Expanded age ranges, careful consideration of covariates that modify the rate of follicular depletion or other aging processes, tracking of women over long periods of time, and plans for specimen and data storage and sharing should all be considered. Although there are limits to what any single study can do, collaborative efforts across multiple study sites may have potential for identifying key linkages between reproductive aging and age- and disease-related health issues. In addition, future studies will likely be enhanced by ever-improving technological innovations in the areas of biological imaging, hormonal assays, blood sampling methods, genetic methods, and bioinformatics.

Acknowledgments

Dr. Ferrell contributed to this article in her personal capacity. The views expressed are her own and do not necessarily represent the views of the National Institutes of Health or the United States Government.

Conflicts of interest

The authors declare no conflicts of interest.

References