

Ovarian follicle loss in humans and mice: lessons from statistical model comparison

J.E. Coxworth* and K. Hawkes

Department of Anthropology, University of Utah, Salt Lake City, UT 84112, USA

*Correspondence address. E-mail: james.coxworth@anthro.utah.edu

Submitted on March 26, 2010; resubmitted on May 3, 2010; accepted on May 7, 2010

BACKGROUND: Mammalian oocyte stocks reach maximum size in early development and begin depletion immediately thereafter. This depletion ends women's fertility by midlife. Here we compare five models proposed to characterize human follicular depletion, highlight underlying variation in atresia, and use oocyte counts from laboratory mice to illustrate possible effects of known covariates.

METHODS: We compared statistical models, of human data, from five well-known sources and also compared the models' fit to data from four genetically distinct strains of mice.

RESULTS: A model first published by Hansen *et al.* (2008) fit the human data better than any of the alternatives. Best-fit models of oocyte loss in the four strains of mice differed substantially from the best-fit model of the aggregated mouse data.

CONCLUSIONS: Although the power model published by Hansen *et al.* (2008) fit the human data best, Faddy and Gosden's (1996) differential equation model may be a more useful characterization of human follicular atresia. However, these models leave a great deal of variation unexplained. Mouse strain comparisons show that follicle loss in genetically distinct subpopulations can differ substantially from the pattern in the aggregate population. This indicates that differences in follicular stock size between and within populations depend upon more than a single predictor (i.e. age or follicle stocks at previous time points). Our reliance upon data from Western populations represents this study's most important limitation. Expanding data collection to include likely covariates and a wider range of human populations would improve the basis for predicting individual trajectories of follicle loss as more women worldwide opt to delay childbearing and risk aging beyond their own windows of fertility.

Key words: Akaike information criteria / atresia / biphasic / menopause / oocyte

Introduction

The challenge of characterizing the human pattern of ovarian follicle loss has inspired a number of statistical models over the past three decades (see Mattison, 1985; Thomford *et al.*, 1987; Faddy *et al.*, 1992; Faddy and Gosden, 1996; Faddy, 2000; Hansen *et al.*, 2008). Though difficult, modeling follicle loss in the human ovary is important for a number of reasons. We are concerned with two. First, as evolutionary anthropologists, we seek an accurate characterization of human follicle loss to help understand distinctive features of human life history (e.g. Cant and Johnstone, 2008). Statistical modeling can help distinguish derived from conserved features in the evolution of the human lineage by facilitating cross-species comparisons of follicular atresia (Robson *et al.*, 2006; Jones *et al.*, 2007; Hawkes, 2006; Hawkes *et al.*, 2009; Hawkes and Smith, 2010). Second, models of follicular atresia might help women make informed reproductive decisions (Faddy and Gosden, 1996; Broekmans *et al.*, 2009). But, as others

have recognized, the current models have limited utility for that job. The model comparisons we have conducted demonstrate the pitfalls of forecasting individual reproductive trajectories using extant models.

Oocyte stocks are established before birth in humans (Baker, 1963) and then begin to decline, mostly through atresia, at rates very similar to those described in chimpanzees (Jones *et al.*, 2007). While chimpanzees usually die before the end of their childbearing years, women do not (Hawkes, 2003; Hawkes and Blurton Jones, 2005). Even in human populations with life expectancies less than 40, girls who survive childhood usually outlive their fertility (Howell, 1979; Hill and Hurtado, 1996; Blurton Jones *et al.*, 2002). This difference between ovarian and somatic aging is a distinctive feature of human life history (Robson *et al.*, 2006). Human post-menopausal longevity reveals individual differences in ovarian aging that are obscured by death in our closest living relatives (Hawkes and Smith, 2010). Without an accurate model of human follicle loss we cannot conclusively discriminate unique attributes of reproductive aging in our species.

In addition to these evolutionary issues, the character and cause of individual differences in initial follicular stocks and trajectories of depletion are of growing interest now as more women delay child-bearing to ages at which natural pregnancies are unlikely (Broekmans *et al.*, 2009). Previous research suggests that an accurate description of follicular atresia could help forecast the timing of important thresholds in declining fertility and so alert women to time remaining on their biological clocks (Faddy and Gosden, 1996).

Over the past two decades, many statistical models have been advanced to account for changes in follicle numbers across the human lifespan (Mattison, 1985; Thomford *et al.*, 1987; Faddy *et al.*, 1992; Gougeon *et al.*, 1994; Faddy and Gosden, 1996; Faddy, 2000; Hansen *et al.*, 2008). During that time, one model—which we refer to as the biphasic model—developed by Faddy *et al.* (1992) became the conventional wisdom for many researchers (e.g. Fitzgerald *et al.*, 1998; Kline *et al.*, 2000; Al-Azzawi, 2001; Lobo, 2005; Cant and Johnstone, 2008; Tilly and Telfer, 2009) in spite of telling criticism (Leidy *et al.*, 1998), even by its authors (Faddy and Gosden, 1996). A consensus on which model is the most helpful has not been reached, however, and researchers continue to propose potentially useful alternatives.

These models describe mean trajectories of follicle loss and allow comparisons between populations and species, but they leave substantial within-group variation unexplained. The magnitude of that variation makes them inadequate guides for personal planning (te Velde and Pearson, 2002; Broekmans *et al.*, 2009; Hawkes *et al.*, 2009). Hansen *et al.* (2008) emphasized this problem as they introduced a new method for counting non-growing follicles, a new data set, and a new model of follicular atresia. They note a 25-fold difference in primordial follicle numbers in two of their 30-year-old subjects, making their model (which we consider below) ‘inadequate for predicting the reproductive lifespan for an individual’ (p. 206). They suggest that, ‘the incorporation of other variables such as clinical markers of ovarian reserve (e.g. early follicular phase FSH and antral follicle counts obtained by transvaginal ultrasound examination) and lifestyle factors would improve the predictive power of the model’ (p. 207). Some of the variation in follicle counts by age could be random (Finch and Kirkwood, 2000), but a portion is demonstrably linked to quantifiable traits (Westhoff *et al.*, 2000; Hardy and Kuh 2002; Broekmans *et al.*, 2009; Tom *et al.*, 2010). Smoking is one frequently nominated example (Do *et al.*, 1998; Kato *et al.*, 1998; Harlow and Signorello, 2000; Westhoff *et al.*, 2000; Gold *et al.*, 2001; Palmer *et al.*, 2003; Kinney *et al.*, 2007).

In addition to heterogeneity within populations, distributions appear to differ between human populations as well (Thomas *et al.*, 2001; Sievert, 2006; Bentley and Muttukrishna, 2007). Some of this variation is surely due to methodological discrepancies. Age at menopause varies depending on measurement techniques (Sievert and Hautaniemi, 2003) and inter-population differences are reduced when methods are held constant (Morabia and Costanza, 1998). But many careful studies (e.g. Goodman *et al.*, 1985; Sarin *et al.*, 1985; Wood *et al.*, 1985; Bentley and Muttukrishna, 2007) find mean ages at menopause that depart from the classic Minnesota study of Treloar *et al.* (1981). An analysis of ages at menopause for different racial/ethnic subsets of American women showed that menopausal age covaries with group identification (Henderson *et al.*, 2008). This finding suggests that the current data set, drawn from a subset of Western (specifically, American and European) populations, probably does not represent the full range of variation in human follicle loss.

Surprisingly, inbred laboratory mice show levels of individual variation similar to that observed in the classic human samples. As noted by Finch and Kirkwood (2000: p. 23), ‘there is a threefold range in the numbers of ovarian oocytes present at birth among individuals in inbred mice as well as in unrelated humans’. And the variation is even greater in adults, ‘at the onset of reproductive senescence in mice, oocyte numbers vary from none . . . up to 1000 in others of the same age. . . studies of premenopausal women. . . showed a 1000-fold range of oocytes’ (Finch and Kirkwood 2000: p. 80). To demonstrate the effects of measurable traits on follicle loss at the level of the subpopulation, we fit models of follicular atresia to oocyte counts from four genetically distinct strains of laboratory mice (Jones and Krohn, 1961; Faddy *et al.*, 1976, 1983) and then compared the best-fit models of each strain to one another.

Our analyses consisted of statistical model comparisons conducted with two data sets. In order to better characterize the pattern of follicle loss in humans, facilitate cross-species comparisons, and define the range of variation in follicular atresia, we fitted five unique models to follicle counts from 238 human females and then compared them using Akaike Information Criteria or AIC (Akaike, 1974). As mentioned above, we also fit the three most flexible, though not the most biologically plausible, models to mouse oocyte counts in order to highlight similarities and differences between strains. These comparisons produced two main findings. First, models presented by Hansen *et al.* (2008) and Faddy and Gosden (1996; Faddy, 2000) do the best job of describing human follicle loss in the current sample. Second, population-wide models of follicular atresia can differ dramatically from the best-fit models for subsets of that population.

Materials and Methods

Data

Human data came from five sources: Block (1952, 1953), Forabosco and Sforza (2007), Gougeon *et al.* (1994) and Richardson *et al.* (1987), and the recent sample published by Hansen *et al.* (2008). Detailed descriptions of these data can be found in the original publications. We included all data from all sources except Forabosco and Sforza (2007). Since it remains unclear when prenatal follicle stocks stop growing (Forabosco and Sforza, 2007: p. 681), we have elected to include follicular data gathered nine months or more after conception. To the best of our knowledge, this data set constitutes the largest aggregation of human follicular data analyzed thus far ($n = 238$).

The mouse data came from Jones and Krohn (1961). They reported oocyte numbers and ages for parous and nulliparous mice from four genetic strains ($n = 275$). As with human data, a more detailed description of the data and their collection can be found in the source article. Since the authors did not report oocyte numbers and ages for each mouse, one of us (J.E.C.; at the recommendation of Roger Gosden) obtained approximate values by measuring the location of data points presented graphically and calculating the oocyte number and age in days based on the points’ locations. Since we were not concerned with variation in oocyte counts with parity, we pooled the data from parous and nulliparous mice prior to analysis.

Models

We considered five unique models to approximate mean follicle loss. Four of the five functions have been used previously. We included one

unpublished model based on the Gompertz function in our comparisons. Initially, we modeled both changes in the mean number of follicles with age and changes in the variance. However, this process introduced a number of complications and contributed little to our understanding of follicle loss. As a consequence, we used a single, constant variance parameter (v). In all functions except the differential equation (Eqs. (5–7)), y symbolizes follicle counts and x is age.

The simplest functional form we included in model comparisons was exponential, similar models published by [Thomford et al. \(1987\)](#) and [Matison \(1985\)](#). The parameters to be estimated are labeled a and b :

$$y = \exp(a + bx) \quad (1)$$

We also used a biphasic or bi-exponential model similar to the version first published by [Faddy et al. \(1992\)](#):

$$y = \exp(a_1 + b_1 x), \text{ for } y < y_c \quad (2)$$

$$y = \exp(a_2 + b_2 x), \text{ for } y > y_c \quad (3)$$

$$y_c = \exp\left(\frac{-b_1 a_2 + b_2 a_1}{b_2 - b_1}\right) \quad (4)$$

In this model, a_1 , a_2 , b_1 and b_2 are the parameters to be estimated.

We also considered a differential equation model ([Faddy and Gosden, 1996](#); [Faddy, 2000](#)), where the number of follicles at each point (i) depends upon the number of follicles at the previous point ($i - 1$). This model generalizes the idea described by [Faddy et al. \(1992\)](#) that follicle loss transitions from a low initial rate to a higher rate as follicle stocks decline. The equation at the root of this model is:

$$\frac{dy}{dx} = -\left(\alpha + \frac{\beta}{\gamma + y}\right)y \quad (5)$$

The solution is the limit of y_i as i approaches infinity. Values of y_i are given by the following equations:

$$y_1 = y_0 \exp\left(-\frac{\alpha\gamma + \beta}{\gamma}x\right) \quad (6)$$

$$y_i = y_{i-1} - \frac{(\beta/\alpha) \log[(\alpha y_{i-1} + \alpha\gamma + \beta)/(\alpha\gamma_0 + \alpha\gamma + \beta)] + \gamma \log(y_{i-1}/y_0) + (\alpha\gamma + \beta)x}{\beta/(\alpha y_{i-1} + \alpha\gamma + \beta) + \gamma/y_{i-1}}, \text{ for } i \geq 2 \quad (7)$$

The parameters to be estimated are α , β and γ . We drew the appropriate value for y_0 (i.e. mean number of follicles at time zero) from the data.

The final published model we incorporated was the power function proposed by [Hansen et al. \(2008\)](#):

$$y = \exp(ax^b + c) \quad (8)$$

There are three parameters to be estimated in this model: a , b and c . As mentioned above, we also included a previously unused model based on the Gompertz function:

$$y = a(1 - \exp(-b \exp(-cx))) \quad (9)$$

As with the power function, there are three parameters to estimate for this model: a , b and c . [Gougeon et al.'s \(1994\)](#) model III would seem an obvious candidate for inclusion in model comparisons since it describes a progressive decline in follicle stocks with age. It incorporates five parameters, however, which makes it more complex than the power function model, while less biologically plausible than the differential equation model. As a consequence, we elected to exclude it.

Statistical analyses

To allow for straightforward comparisons between our results and those of previous researchers ([Faddy et al., 1992](#); [Gougeon et al., 1994](#); [Faddy, 2000](#); [Hansen et al., 2008](#)), we log-transformed human follicle and mouse oocyte counts before model fitting. [Leidy et al. \(1998\)](#) noted that log-transforming follicle counts erroneously increases the appeal of a biphasic model since the transformed data display a sharp bend when plotted against age. Yet, log transformation also highlights deviations from a constant rate of loss, a useful attribute when studying follicular atresia. Since both data sets included zeroes, we add a small constant (0.1) to all follicle counts before log transformation. We assumed normally distributed error in log-transformed follicle counts; an assumption justified by [Faddy et al. \(1992\)](#). Using the `mle2` function in R statistical package ([R Development Core Team, 2008](#)), we estimated parameters via maximum likelihood.

After parameter estimation, we compared the relative fit of each model of the mean using AIC ([Akaike, 1974](#)). AIC compares the relative predictive power of statistical models (see [Anderson et al., 2000](#)). AICc is a modification of AIC specially formulated for analyses where the ratio of data points to parameters is relatively small. Although not all of our models required the use of this modification, AICc produces the same results as AIC with increased sample size to parameter ratios ([Burnham and Anderson, 2004](#)). We therefore used AICc in all model comparisons for the sake of consistency. We relied most heavily upon AICc weights in comparing models. As described by [Anderson et al. \(2000:918\)](#), AICc weights provide a relative ranking of statistical models' predictive value by quantifying the informational evidence in favor of each. Whereas AICc scores provide an absolute index of each model, with lower scores leading to a higher rank, AICc weights are relative: each model's predictive utility is ranked against all other models in the set. Higher-ranked models have larger AICc weights. All else being equal, models with fewer parameters receive higher ranks. As a consequence, competing models may have similar AICc scores and very different weights. We report both the absolute AICc scores and relative AICc weights below.

Results

Human model comparison

In the human data set, the power model ([Hansen et al., 2008](#)) out-ranked all others (AICc = 616.4, weight = 0.55; Table I and Fig. 1), with the differential equation model ([Faddy and Gosden, 1996](#); [Faddy, 2000](#)) ranked a close second (AIC = 616.8, weight = 0.43; Table I and Fig. 1). The biphasic model ([Faddy et al., 1992](#)) received an AICc weight slightly greater than zero (AICc = 623.0, weight = 0.02; Table I and Fig. 2). The Gompertz model (AICc = 626.9, weight = 0.00; Table I and Fig. 2) and the exponential model were ranked last (AICc = 662.2, weight = 0.00; Table I and Fig. 2). Judging from both absolute AICc scores and AICc weights, the power model and differential equation model appear to be superior descriptions of human follicular atresia.

Mouse model comparison

We compared models of oocyte loss for each genetic strain of mice, as well as for the aggregated population (i.e. data from all strains lumped together). The biphasic model was top-ranked for the aggregated data set (AICc = 518.5, weight = 0.49; see Table II and Fig. 3). The exponential model followed closely behind (AICc = 519.2, weight = 0.34). The power model ranked last (AICc = 520.7, weight = 0.17). Model comparisons conducted with data from

Table 1 Models of human follicle loss.

Model	Mean	AICc	DF	Weight	Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
Power	Eq. (8)	616.4	4	0.55	$a = -0.00008$ (-0.002, -0.00005)	$b = 3.30$ (2.46, 4.43)	$c = 5.56$ (5.30, 5.83)	$v = 0.87$ (0.80, 0.95)	—
Differential	Eqs (6 and 7)	616.8	4	0.43	$\alpha = -2.52$ (-2.64, -2.44)	$\beta = 952.06$ (706.21, 1279.16)	$\gamma = 664.15$ (384.56, 1616.09)	$v = 0.87$ (0.80, 0.95)	—
Biphasic	Eqs (2 and 3)	623.0	5	0.02	$a_1 = 5.85$ (5.49, 6.17)	$a_2 = 10.70$ (9.56, 12.15)	$b_1 = -0.03$ (-0.05, -0.01)	$b_2 = -0.16$ (-0.19, -0.13)	$v = 0.88$ (0.80, 0.96)
Gompertz	Eq. (9)	626.9	4	0.00	$a = 302625.60$ (172726.10, 549006.40)	$b = 3092.47$ (268.19, 127014.50)	$c = 0.29$ (0.23, 0.36)	$v = 0.88$ (0.81, 0.97)	—
Exponential	Eq. (1)	662.2	3	0.00	$a = 6.36$ (6.05, 6.66)	$b = -0.06$ (-0.07, -0.06)	$v = 0.96$ (0.88, 1.05)	—	—

Values in parentheses indicate 95% confidence intervals.

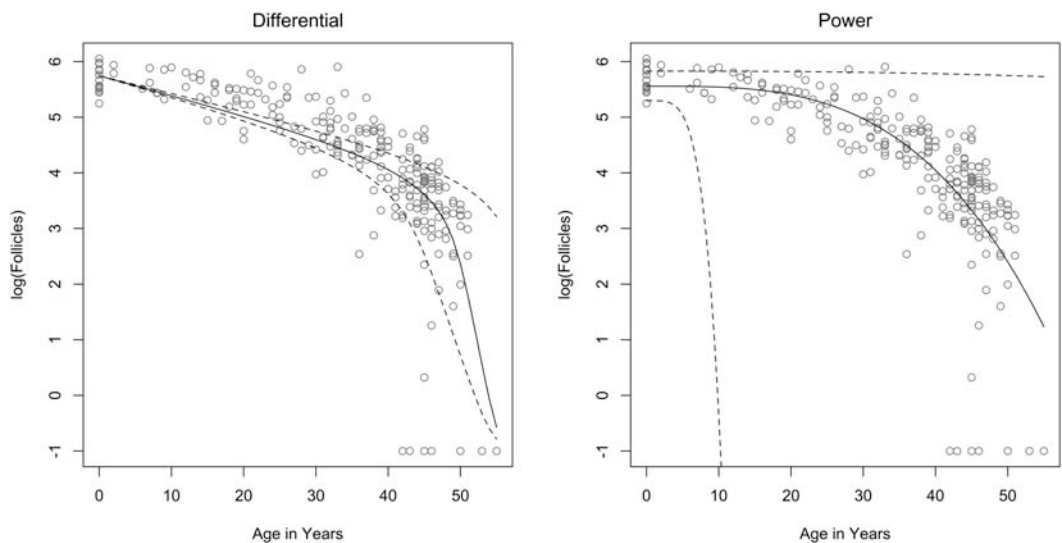


Figure 1 Highest ranked models of human follicle loss (plot shows log of follicle counts versus age in years). Solid line is the fitted model and dashed lines represent 95% confidence intervals ($n = 238$).

individual strains of mice generally mirrored this ranking, but with the difference that the biphasic model was even more highly ranked (Strain A: AICc = 77.7, weight = 0.91; Strain CBA: AICc = 149.0, weight = 1.0; Strain CBAxA: AICc = -78.7, weight = 0.83; Strain RIII: AICc = -61.3, weight = 1.0; see Table III and Fig. 4).

Discussion

Our findings, in concert with research on physiological mechanisms of follicle loss, add to the reasons to reject the biphasic model of human follicle loss of Faddy et al. (1992), suggesting as it does a population-wide physiological shift in women as they reach the age of 37.5 (Faddy and Gosden, 1996; Leidy et al., 1998; Hawkes and Smith, 2010). Though the power model (Hansen et al., 2008) was ranked

higher, the differential equation model developed by Faddy (Faddy and Gosden, 1996; Faddy, 2000) may be a more useful descriptor of human follicular atresia because it accords with recent findings regarding the physiology of follicle loss (see discussion below; Da Silva-Buttkus et al., 2009). However, wide variation around mean follicle count at most ages highlights the inadequacy of any model of loss based upon a single input variable (i.e. age or previous follicle counts) for predicting individual trajectories.

Our analysis of the follicle counts in inbred strains of mice revealed subpopulation differences. Although ‘the variations within a [mouse] strain approximate the variations between strains in the average oocyte number and in the timing of reproductive senescence’ (Finch, 2007: p. 308), the differences are not trivial. Our mouse comparisons indicate two aspects of atresia that likely characterize humans

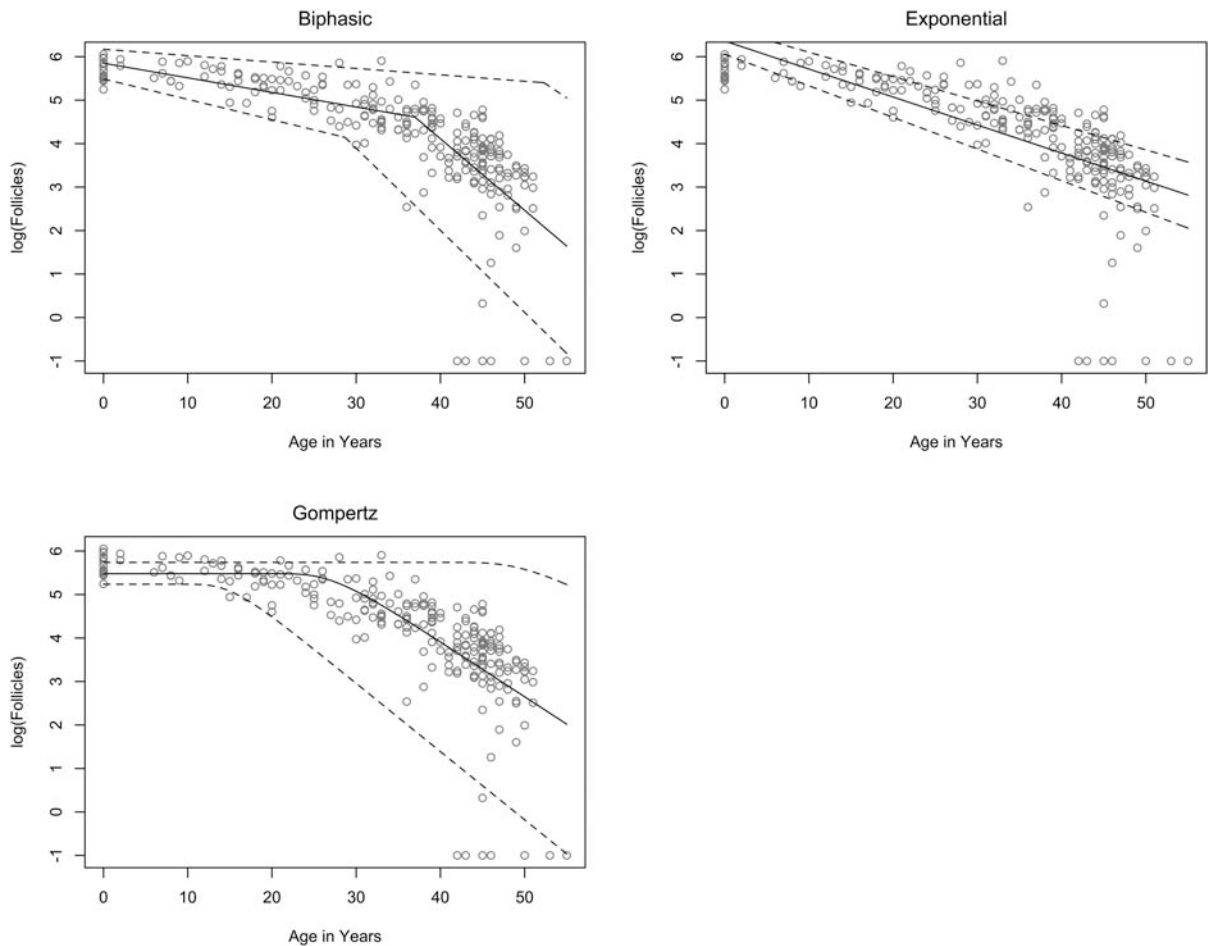


Figure 2 Lower ranked models of human follicle loss (plot shows log of follicle counts versus age in years). Solid lines are fitted models and dashed lines represent 95% confidence intervals ($n = 238$).

Table II Models of oocyte loss in aggregated mouse data (all strains).

Model	Mean	AICc	DF	Weight	Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
Biphase	Eqs (2 and 3)	518.5	5	0.49	$a_1 = 3.93$ (3.83, 4.03)	$a_2 = 16.99$ (6.22, 18.42)	$b_1 = -0.004$ (-0.004, 0.003)	$b_2 = -0.02$ (-0.02, -0.02)	$v = 0.61$ (0.56, 0.66)
Exponential	Eq. (1)	519.2	3	0.34	$a = 3.94$ (3.84, 4.05)	$b = -0.004$ (-0.004, -0.004)	$v = 0.62$ (0.57, 0.67)	—	—
Power	Eq. (8)	520.7	4	0.17	$a = -0.002$ (-0.008, -0.001)	$b = 1.08$ (0.88, 1.26)	$c = 3.90$ (3.77, 4.05)	$v = 0.61$ (0.57, 0.67)	—

Values in parentheses indicate 95% confidence intervals.

as well. First, genetic differences between subpopulations can have substantial effects on trajectories of oocyte loss. Second, genetic differences alone do not account for all of the variation observed within strains. Recent research suggests that stochastic and environmental factors, both pre- and postnatal, can have measurable effects on ovarian aging (Ibanez *et al.*, 2000, 2003; Hardy and Kuh, 2002;

Gosden *et al.*, 2007; Tom *et al.*, 2010). Our results agree with these earlier findings.

Population versus individual trajectories

As shown in Fig. 4 and Table III, mean follicle loss follows different trajectories in the four strains of mice studied by Jones and Krohn (1961;

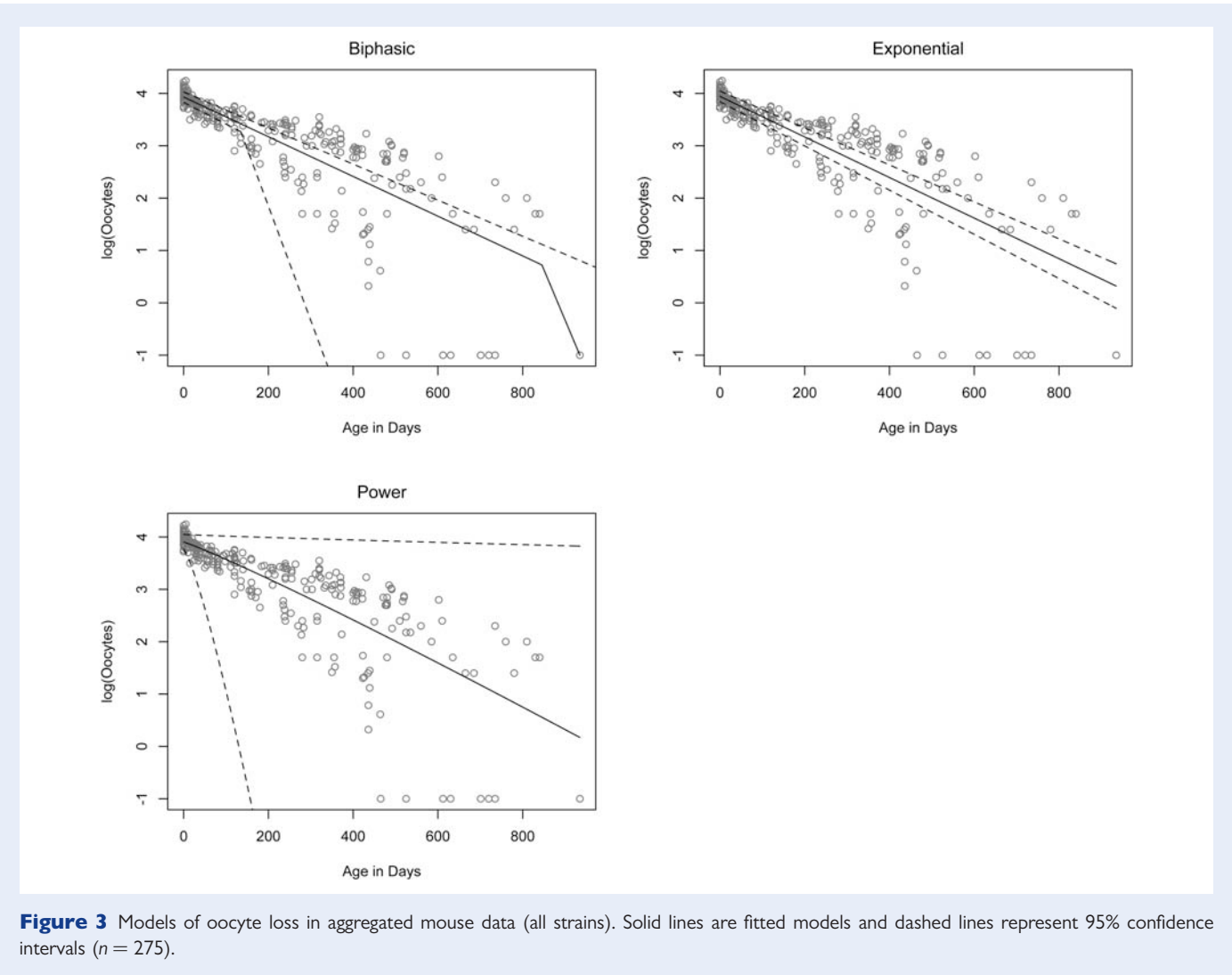


Table III Description of highest ranked models of oocyte loss in each mouse strain.									
Strain	Mean	AICc	DF	Weight	Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
A	Eqs (2 and 3)	77.7	5	0.91	$a_1 = 3.89$ (3.76, 4.02)	$a_2 = 6.58$ (4.78, 7.22)	$b_1 = -0.003$ (-0.003, -0.002)	$b_2 = -0.009$ (-0.01, -0.008)	$v = 0.38$ (0.33, 0.46)
CBA	Eqs (2 and 3)	149.0	5	1.00	$a_1 = 3.95$ (3.78, 4.12)	$a_2 = -1.39$ (-2.94, 0.54)	$b_1 = -0.007$ (-0.007, -0.006)	$b_2 = 0.002$ (0.0001, 0.005)	$v = 0.53$ (0.46, 0.62)
CBAxA	Eqs (2 and 3)	-78.7	5	0.83	$a_1 = 3.95$ (3.90, 3.99)	$a_2 = 4.26$ (4.09, 4.64)	$b_1 = -0.002$ (-0.002, -0.002)	$b_2 = -0.003$ (-0.003, -0.002)	$v = 0.11$ (0.09, 0.13)
RIII	Eqs (2 and 3)	-61.3	5	1.00	$a_1 = 3.92$ (3.87, 3.97)	$a_2 = 6.25$ (5.96, 6.77)	$b_1 = -0.002$ (-0.003, -0.002)	$b_2 = -0.007$ (-0.007, -0.006)	$v = 0.13$ (0.11, 0.16)

The biphasic model was ranked highest for all strains.

see also Faddy et al., 1976, 1983). This result was somewhat surprising, since the strains were derived from a highly inbred population of laboratory mice (Jones and Krohn, 1961) with relatively little genetic diversity. Differences in the shape of atresia led to highly inconsistent predictions of oocyte numbers. For example, predicting the number of remaining oocytes for a 200 day-old mouse with the aggregate population model would yield an estimate of 1349. This number conflicts with the estimate of 1950 oocytes using the A strain model, 355 oocytes using the CBA model, 3548 oocytes using the CBAxA model and 3311 oocytes using the RIII model. Genetic differences may have an equally important role in determining ovarian aging in humans (Voorhuis et al., 2010).

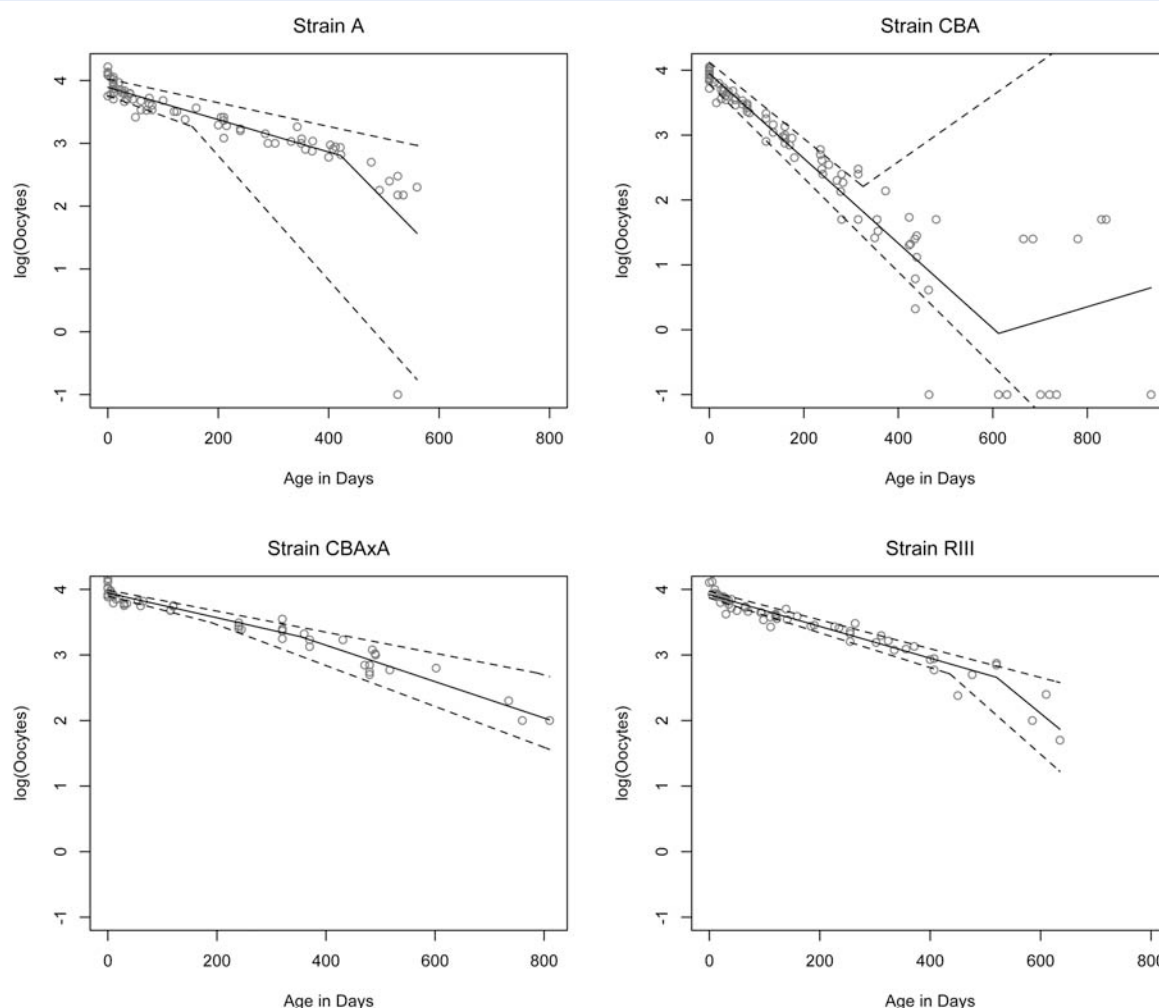


Figure 4 Highest-ranked models of oocyte loss in each mouse strain. The biphasic model was ranked highest for all strains. Solid lines are fitted models and dashed lines represent 95% confidence intervals ($n = 275$).

Since genetic diversity within the strains is low, much of the within-strain variation must be either due to chance or the influence of environmental factors on initial oocyte stocks and rates of loss. Treloar *et al.* (2000) reported substantial variation in the timing of menopause between monozygotic human twins—attesting to differences in ovarian aging among genetically identical individuals in our own species. While there are differences between laboratory rodents and humans in oocyte development and loss (McGee and Hsueh, 2000; Adhikari and Liu, 2009), diversity within mouse strains could result from the kinds of variation in pre- and post-natal somatic growth and development that have been nominated as important determinants of human oocyte stocks and rates of loss (Ibanez *et al.*, 2000, 2003; Hardy and Kuh, 2002; Gosden *et al.*, 2007; Hart *et al.*, 2009; Tom *et al.*, 2010).

Variation in rates of oocyte loss among mouse strains leads us to expect a range of trajectories in human subpopulations and, as a consequence, similarly inaccurate predictions when relying upon a single model based solely on age or stock size in the previous time interval. Forecasting a woman's age at menopause with such models is likely to

be a less accurate process than previous research suggests (see Faddy and Gosden, 1996; Lobo, 2005). Longitudinal samples would be the ideal way to investigate individual trajectories but primordial or non-growing follicles can only be counted in surgically removed or autopsied ovaries. As a consequence, samples are necessarily cross-sectional. Expanding these samples to include likely covariates (e.g. smoking, birthweight, weight in infancy, ethnic/racial background, socioeconomic standing, early follicular phase FSH, antral follicle counts, inhibin-B and AMH recorded prior to surgery) and individuals from a more diverse range of human populations and subpopulations is a feasible way to improve cross-sectional data.

Modeling human follicle loss

In spite of longstanding critiques (Faddy and Gosden, 1996; Leidy *et al.*, 1998), the biphasic model is still used to describe human follicle loss (Fitzgerald *et al.*, 1998; Kline *et al.*, 2000; Al-Azzawi, 2001; Lobo, 2005; Cant and Johnstone, 2008; Tilly and Telfer, 2009). Our results add to previous objections. Statistical measures of predictive utility rank it far below the differential equation model and power model

(Table I). As authors of the biphasic model noted themselves (Faddy and Gosden, 1996), it implies a biologically implausible process. Both the power model and the differential equation model are clearly preferable to the widely cited biphasic model. Of these two, the differential equation model (Faddy and Gosden, 1996; Faddy, 2000) is more consistent with recent developments in the study of mechanisms of follicle depletion.

Research into these physiological mechanisms increasingly points to intra-ovarian communication as an influential factor in the recruitment of primordial follicles. The notion that follicular loss depends upon the number of remaining follicles formed the foundation of Faddy and Gosden's (1996) introduction of the differential equation model. They based their idea upon an uncertain supposition that paracrine factors might play a role in atresia (Faddy and Gosden, 1996: p. 1486). Their supposition has gained substantial support in the intervening years. Investigators working with reproductive specimens from a variety of mammals have postulated that resting primordial follicles depend upon inhibitory signals to avoid initial recruitment, though it remains unclear where such signals originate (for reviews of relevant research see McGee and Hsueh, 2000; Adhikari and Liu, 2009). Nilsson *et al.* (2007) showed that anti-Müllerian hormone secreted by granulosa cells surrounding primary and secondary follicles inhibits follicle activation in rodents (see also Visser and Themmen, 2005). Recent findings reported by Da Silva-Buttkus *et al.* (2009) suggest that primordial follicles may themselves produce the inhibitory signal suggested by previous researchers. Though the specific hormonal agents responsible for inhibition remain unidentified, these findings support the premise of the differential equation model that a follicle's chance of undergoing initial recruitment depends upon the number of follicles remaining in the ovary.

Limitations

Although biphasic models fit the mouse data much better than the alternatives, they are no more biologically plausible for mice than humans. Inspection suggests that fitting a biologically plausible model to the mouse data will not be easy. We did not try to construct one. The biphasic model, as well as the power and exponential models, serve here for their simplicity and flexibility and not for their utility as descriptions of physiological processes. We used them to demonstrate differences among identifiable subpopulations and to underscore the variation that remains even when genetic differences are minimal. We look forward to the development of biologically informed models of follicle loss in rodents.

The differential equation model is the best description of human follicle loss for the current human sample, but it remains unclear how it would fit subpopulations distinguished by initial stock size (if that could be assayed) or other covariates. The data used in model fitting come from Western Caucasian women for whom early developmental characteristics and other likely covariates are not available. Potential differences in initial stocks and rates of atresia in non-Western women remain unknown, as does the range of variation among more diverse groups of Westerners. Quantitative studies of the timing of menopause show a range of variation between populations (Morabia and Costanza, 1998; Thomas *et al.*, 2001; Bentley and Muttukrishna, 2007). This likely relates to ovarian follicle counts by age (Westhoff *et al.*, 2000). Follicle counts that include covariates

and come from a wider range of human groups would allow investigation of inter- and intra-populational differences. The medical and social importance of variation between and within groups will continue to increase as ages at first birth climb higher and fertility reaches evolutionarily unprecedented lows.

Authors' roles

J.E.C. conducted all quantitative analyses, compiled the data and drafted this paper. K.H. helped compile the data and provided assistance in drafting this paper.

Acknowledgements

We thank Malcolm Faddy for comments on previous drafts of this paper and assistance in implementing his differential equation model. We appreciate Roger Gosden's guidance in recovering the data reported in Jones and Krohn (1961) and Alain Gougeon's correspondence regarding the origins of his and Block's data sets. Finally, we thank Richard McElreath for sharing both his statistical and computer programming expertise.

Funding

This work was supported by National Science Foundation grant 0850951 (Chimpanzee Reproductive and Physiological Aging).

References

- Adhikari D, Liu K. Molecular mechanisms underlying the activation of mammalian primordial follicles. *Endocr Rev* 2009;**30**:438–464.
- Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr* 1974;**19**:716–723.
- Al-Azzawi F. The menopause and its treatment in perspective. *Postgrad Med J* 2001;**77**:292–304.
- Anderson DR, Burnham KP, Thompson WL. Null hypothesis testing: problems, prevalence, and an alternative. *J Wildlife Manage* 2000;**64**:912–923.
- Baker TG. A quantitative and cytological study of germ cells in human ovaries. *Proc R Soc Lond B* 1963;**158**:417–433.
- Bentley GR, Muttukrishna S. Potential use of biomarkers for analyzing interpopulation and cross-cultural variability in reproductive aging. *Menopause* 2007;**14**:668–679.
- Block E. Quantitative morphological investigations of the follicular system in women. Variations at different ages. *Acta Anat* 1952;**14**:108–123.
- Block E. A quantitative morphological investigation of the follicular system in newborn female infants. *Acta Anat* 1953;**17**:184–205.
- Blurton Jones NG, Hawkes K, O'Connell F. Antiquity of postreproductive life: are there modern impacts on hunter-gatherer postreproductive life spans? *Am J Hum Biol* 2002;**14**:184–205.
- Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. *Social Methods Res* 2004;**33**:261–304.
- Broekmans FJ, Soules MR, Fauser BC. Ovarian aging: mechanisms and clinical consequences. *Endocr Rev* 2009;**30**:465–493.
- Cant MA, Johnstone RA. Reproductive conflict and the separation of reproductive generations in humans. *PNAS* 2008;**105**:5332–5336.
- Da Silva-Buttkus P, Marcelli G, Franks S, Stark J, Hardy K. Inferring biological mechanisms from spatial analysis: prediction of a local inhibitor in the ovary. *Proc Natl Acad Sci* 2009;**106**:456–461.

- Do KA, Treloar SA, Pandeya N, Purdie D, Green AC, Heath AC, Martin NG. Predictive factors of age at menopause in a large Australian twin study. *Hum Biol* 1998;**70**:1073–1091.
- Faddy MJ. Follicle dynamics during ovarian ageing. *Mol Cell Endocrinol* 2000;**163**:43–48.
- Faddy MJ, Gosden RG. A model conforming the decline in follicle numbers to the age of menopause in women. *Hum Reprod* 1996;**11**:1484–1486.
- Faddy MJ, Jones EC, Edwards RG. An analytical model for ovarian follicle dynamics. *J Exp Zool* 1976;**197**:173–185.
- Faddy MJ, Jones EC, Edwards RG. Ovarian follicle dynamics in mice: a comparative study of three inbred strains and an F1 hybrid. *J Endocrinol* 1983;**96**:23–33.
- Faddy MJ, Gosden RG, Gougeon A, Richardson SJ, Nelson JF. Accelerated disappearance of ovarian follicles in mid-life: implications for forecasting menopause. *Hum Reprod* 1992;**7**:1342–1346.
- Finch CE. *Biology of Human Longevity: Inflammation, Nutrition, and Aging in the Evolution of Life Spans*. San Diego, CA: Academic, 2007.
- Finch CE, Kirkwood TBL. *Chance, Development, and Aging*. New York: Oxford, 2000.
- Fitzgerald C, Zimon AE, Jones EE. Aging and reproductive potential in women. *Yale J Biol Med* 1998;**71**:367–381.
- Forabosco A, Sforza C. Establishment of ovarian reserve: a quantitative morphometric study of the developing human ovary. *Reprod Endocrinol* 2007;**88**:675–683.
- Gold EB, Bromberger J, Crawford S, Samuels S, Greendale GA, Harlow SD, Skurnick J. Factors associated with age at natural menopause in a multiethnic sample of midlife women. *Am J Epidemiol* 2001;**153**:865–874.
- Goodman MJ, Estioko-Griffin A, Griffin PB, Grove JS. Menarche, pregnancy, birth spacing and menopause among the Agta foragers of Cagayan Province, Luzon, the Philippines. *Ann Hum Biol* 1985;**12**:169–177.
- Gosden RG, Treloar SA, Martin NG, Cherkas LF, Spector TD, Faddy MJ, Silber SJ. Prevalence of premature ovarian failure in monozygotic and dizygotic twins. *Hum Reprod* 2007;**22**:610–615.
- Gougeon A, Ecochard R, Thalabard JC. Age-related changes of the population of human ovarian follicles: increase in the disappearance rate of non-growing and early-growing follicles in aging women. *Biol Reprod* 1994;**50**:653–663.
- Hansen KR, Knowlton NS, Thyer AC, Charleston JS, Soules MR, Klein NA. A new model of reproductive aging: the decline in ovarian non-growing follicle number from birth to menopause. *Hum Reprod* 2008;**23**:699–708.
- Hardy R, Kuh D. Does early growth influence timing of the menopause? Evidence from a British birth cohort. *Hum Reprod* 2002;**17**:2474–2479.
- Harlow BL, Signorello LB. Factors associated with early menopause. *Maturitas* 2000;**35**:3–9.
- Hart R, Slodoba DM, Doherty DA, Norman RJ, Atkinson HC, Newnham JP, Dickinson JE, Hickey M. Prenatal determinants of uterine volume and ovarian reserve in adolescence. *J Clin Endocrinol Metab* 2009;**94**:4931–4937.
- Hawkes K. Grandmothers and the evolution of human longevity. *Am J Hum Biol* 2003;**15**:380–400.
- Hawkes K. Life history theory and human evolution. In: Hawkes K, Paine R (eds). *The Evolution of Human Life History*. Santa Fe, NM: School of American Research, 2006, 45–93.
- Hawkes K, Blurton Jones NG. Human age structures, paleodemography, and the grandmother hypothesis. In: Volland E, Chasiotis A, Schiefenhoel W (ed.), *Grandmotherhood: The Evolutionary Significance of the Second Half of Female Life*. New Brunswick, NJ: Rutgers, 2005, 118–140.
- Hawkes K, Smith KR. Do women stop early? Similarities in fertility decline between humans and chimpanzees. *Ann N Y Acad Sci* 2010; doi: 10.1111/j.1749-6632.2010.05527.x.
- Hawkes K, Smith KR, Robson SL. Mortality and fertility rates in humans and chimpanzees: how within-species variation complicates cross-species comparisons. *Am J Hum Biol* 2009;**21**:578–586.
- Henderson KD, Bernstein L, Henderson B, Kolonel L, Pike MC. Predictors of the timing of natural menopause in the multiethnic cohort study. *Am J Epidemiol* 2008;**167**:1287–1294.
- Hill K, Hurtado AM. *Aché Life History: The Ecology and Demography of a Foraging People*. New York: Aldine de Gruyter, 1996.
- Howell N. *Demography of the Dobe !Kung*. New York: Aldine de Gruyter, 1979.
- Ibáñez L, Potau N, Enriquez G, de Zegher F. Reduced uterine and ovarian size in adolescent girls born small for gestational age. *Pediatr Res* 2000;**47**:575–577.
- Ibáñez L, Potau N, Enriquez G, Marcos MV, de Zegher F. Hypergonadotrophinaemia with reduced uterine and ovarian size in women born small-for-gestational-age. *Hum Reprod* 2003;**18**:1565–1569.
- Jones EC, Krohn PL. The relationships between age, numbers of oocytes and fertility in virgin and multiparous mice. *J Endocrinol* 1961;**21**:469–494.
- Jones KP, Walker LC, Anderson D, Lacreuse A, Robson SL, Hawkes K. Depletion of ovarian follicles with age in chimpanzees: similarities to humans. *Biol Reprod* 1961;**77**:247–251.
- Kato I, Toniolo P, Akhmedkhanov A, Koenig KA, Shore R, Zeleniuch-Jacquotte A. Prospective study of factors influencing the onset of natural menopause. *J Clin Epidemiol* 1998;**51**:1271–1276.
- Kinney A, Kline J, Kelly A, Reuss ML, Levin B. Smoking, alcohol and caffeine in relation to ovarian age during the reproductive years. *Hum Reprod* 2007;**22**:1175–1185.
- Kline J, Kinney A, Levin B, Warburton D. Trisomic pregnancy and earlier age at menopause. *Am J Hum Genet* 2000;**67**:395–404.
- Leidy LE, Godfrey LR, Sutherland MR. Is follicular atresia biphasic? *Fertil Steril* 1998;**70**:851–859.
- Lobo RA. Potential options for preservation of fertility in women. *N Engl J Med* 2005;**353**:64–73.
- Mattison DR. Clinical manifestations of ovarian toxicity. In: RL Dixon (ed.) *Reproductive Toxicology*. New York: Raven Press, 1985. 109–130.
- McGee EA, Hsueh AJW. Initial and cyclic recruitment of ovarian follicles. *Endocr Rev* 2000;**21**:200–214.
- Morabia A, Costanza MC. International variability in ages at menarche, first livebirth, and menopause. *Am J Epidemiol* 1998;**148**:1195–1205.
- Nilsson E, Rogers N, Skinner MK. Actions of anti-Müllerian hormone on the ovarian transcriptome to inhibit primordial to primary follicle 658 transition. *Reproduction* 2007;**134**:209–221.
- Palmer JR, Rosenberg L, Wise LA, Horton NJ, Adams-Campbell LL. Onset of natural menopause in African American women. *Am Public Health Assoc* 2003;**93**:299–306.
- R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing, 2008. ISBN 3-900051-07-0. <http://www.R-project.org>. Austria.
- Richardson SJ, Senikas V, Nelson JF. Follicular depletion during the menopausal transition: evidence for accelerated loss and ultimate exhaustion. *J Clin Endocrinol Metab* 1987;**65**:1231–1237.
- Robson SL, van Schaik CP, Hawkes K. The derived features of human life history. In: Hawkes K, Paine R (eds), *Evolution of Human Life History*. Santa Fe, NM: SAR Press, 2006. 17–44.
- Sarin AR, Singla P, Gupta SK. A 5-year clinicopathological study of 2000 postmenopausal women from Northern India. *Asia-Oceania J Obstet Gynaecol* 1985;**11**:539–544.

- Sievert LL. *Menopause: A Biocultural Perspective*. New Brunswick, NJ: Rutgers, 2006.
- Sievert LL, Hautaniemi SI. Age at menopause in Puebla, Mexico. *Hum Biol* 2003;**75**:205–226.
- te Velde ER, Pearson PL. The variability of female reproductive ageing. *Hum Reprod Update* 2002;**8**:141–154.
- Thomas F, Renaud F, Benefice E, DeMeeus T, Guegan JF. International variability of ages at menarche and menopause: patterns and main determinants. *Hum Biol* 2001;**73**:271–290.
- Thomford PJ, Jelovsek FR, Mattison DR. Effect of oocyte number and rate of atresia on the age at menopause. *Reprod Toxicol* 1987;**1**:41–51.
- Tilly JL, Telfer EE. Purification of germline stem cells from adult mammalian ovaries: a step closer towards control of the female biological clock? *Mol Hum Reprod* 2009;**15**:393–398.
- Tom SE, Cooper R, Kuh D, Guralnik JM, Hardy R, Power C. Fetal environment and early age at natural menopause in a British birth cohort study. *Hum Reprod* 2010;**25**:791–798.
- Treloar AE. Menstrual cyclicity and the pre-menopause. *Maturitas* 1981;**3**:249–264.
- Treloar AE, Sadrzadeh S, Do KA, Martin NG, Lambalk CB. Birth weight and age at menopause in Australian female twin pairs: exploration of the fetal origins hypothesis. *Hum Reprod* 2000;**15**:55–59.
- Visser JA, Themmen AP. Anti-Müllerian hormone and folliculogenesis. *Mol Cell Endocrinol* 2005;**234**:81–86.
- Voorhuis M, Onland-Moret NC, van der Schouw YT, Fauser BCJM, Broekmans FJ. Human studies on genetics of the age at natural menopause: a systematic review. *Hum Reprod* 2010; doi: 10.1093/humupd/dmp055.
- Westhoff C, Murphy P, Heller D. Predictors of ovarian follicle number. *Fertil Steril* 2000;**74**:624–628.
- Wood JW, Johnson PL, Campbell KL. Demographic and endocrinological aspects of low natural fertility in highland New Guinea. *J Biosoc Sci* 1985;**17**:57–79.