Female fertility and longevity

Joshua Mitteldorf

Abstract Does bearing children shorten a woman's life expectancy? Pleiotropic theories of aging predict that it should, and in particular, the Disposable Soma theory predicts unequivocally that this effect should be inescapable. But many demographic studies, historic and current, have found no such effect. In this context, the Caerphilly cohort study stands apart as the sole test that corroborates the theory. Why has this study found an effect that others fail to see? Their analysis is based on Poisson regression, a statistical technique that is accurate only if the underlying data are Poisson-distributed. But the distribution of the number of children born to women in the Caerphilly database departs strongly from Poisson at the high end. This makes the result overly sensitive to a handful of women with 15 children or more who lived before 1700. When these five women are removed from a database of more than 2,900, the Poisson regression no longer shows a significant result. Bilinear regression relating life span to fertility and date of birth results in a small *positive* coefficient for fertility, in agreement with the main trend of reported results.

Keywords Disposable soma · Pleiotropy · Senescence · Aging · Fertility · Life history · Tradeoff

J. Mitteldorf (✉)
Department of Ecology and Evolutionary Biology,
University of Arizona,
Tucson, AZ 85920, USA
e-mail: josh@mathforum.org

Introduction

Evolutionary theory offers compelling reasons to believe that an individual's fertility should be inversely related to life span. This "cost of reproduction" is predicted by all the pleiotropic theories which have dominated the field of life history evolution for more than 50 years. According to these theories, aging has evolved despite its negative effect on individual fitness because of tradeoffs between fertility and longevity. There is powerful selection for alleles that offer fertility enhancement, and some of these have side effects that permit long-term degradation of the soma.

In the original formulation of Williams (1957) of Antagonistic Pleiotropy, the tradeoff was genetic; it is the *potential* for reproduction that has exacted a longevity cost. Alternatively, the Disposable Soma (DS) theory of Kirkwood (1977) is based on a direct metabolic tradeoff between fertility and longevity. Hence, the DS theory in particular predicts there should be a steep and direct cost of reproduction. Within the DS theory, individuals that bear fewer offspring are predicted to live longer than those who bear many offspring, and this remains true regardless of whether curtailment is endemic or contingent. In contrast to the original theory of Williams, the DS theory predicts that animals kept in cages separated by sex and people who use birth control should both have longer life spans than comparable animals and people who are reproducing.

Confirmation of this prediction has been elusive. Some studies showing positive correlations between fertility and longevity, while most find no significant association. Field studies of animals in the wild are difficult, but Ricklefs and Cadena (2007) surveyed 30 species of mammals and birds in captivity, insignificant association, or a slightly positive correlation. Data for humans is abundant but confounded by effects of historical medicine and socioeconomic status. Gavrilova and Gavrilov (2005) provide a comprehensive review of historic studies attempting to link fertility to longevity in human females, going back more than a century to the well-known statistician Karl Pearson (Beeton et al. 1900).

In a large-scale contemporary study of Norwegians, Grundy and Kravdal (2008) found a positive correlation between fertility and longevity. Perls et al. (1997) found a strong positive relationship with late child-bearing in a study of Boston centenarians. Evidence from these and other modern studies has been dismissed without good cause. By the logic of the DS theory, it is the large energy expenditure entailed in pregnancy and lactation that is the root cause of senescence. Benefit from avoidance of this cost should be detectable equally in women whose fertility is curtailed with modern contraception or via abstinence or infertility. Hence, the theory predicts a negative relationship between number of offspring and life span that should be detectable in modern as well as historic demographies. Nevertheless, several researchers have sought to avoid any confounding effect of technological civilization by looking to historical databases.

The DS theory ought also to apply equally to males and to females, but association between male fertility and longevity is also consistently positive (Palmore 1982; Kaplan 1988; Davey Smith et al. 1997), so that some researchers seeking confirmation of the theory have limited their analyses to historic data on human females.

Korpelainen (2000) examined historic data from rural Finns and European aristocrats in the eighteenth and nineteenth centuries. She found positive correlations between number of offspring and life span in both men and women, with the (expected) negative effect only for a small sample of women over 80. Lycett et al (2000) studied a North German historical population in the same time frame and found, again, a "reverse cost" of reproduction, particularly strong among the lower socioeconomic classes. Le Bourg et al (1993) and also Muller et al. (2002) studied French-Canadian populations in this same timeframe and found no evidence for a cost of reproduction (these studies and other demographic data are reviewed by Le Bourg 2001). McArdle et al (2006) studied an Amish population from the eighteenth through the twentieth centuries and found a positive relationship between fertility and longevity. After removing the inverse correlation associated with long-term demographic trends, a positive correlation for men remained, and the correlation for women was insignificant.

The one large-scale human study which claims to discern a negative correlation was published by Kirkwood himself together with J. G. Westendorp (1998). They analyzed women from the Caerphilly database of British aristocracy going back 800 years. Because this publication appeared prominently (in *Nature*) and because it corroborated the theory, it has been widely cited as authority. There are more citations of the Caerphilly study than all the other demographic papers combined (as listed in ISI Web of Science as of October, 2008). For this reason, revisiting the Caerphilly methodology seems worthwhile.

The peerage database

Westendorp and Kirkwood take for their source a published database with family histories of British aristocracy. Their choice of an upper-class sample was made to lower the noise in the data from accidental deaths related to poverty or hardship; historical rather than present data was preferred because the relationship they sought concerned inborn fecundity and could be masked by contraceptive practices that have become common in the twentieth century. Their analysis was limited to women because the reproductive physiology of males is less directly related to the actual number of children sired. The authors further limit analysis to married women for whom there is complete data, who were born before 1875, and who died after menopause (taken alternatively as 50, 55, or 60 years of age); they sought to correlate the number of children each woman bore with her age at death.

The message of their result is announced in the title, "Human longevity at the cost of reproductive

success." Elsewhere, they couch their conclusions in more guarded language. "The relation between age at death and progeny number in British aristocratic women (and men) is consistent with the hypothesis that the longest lived individuals have reduced fertility compared with the majority of the population. This effect is seen most strongly in women born before 1700, for whom the number of children was larger than for women born between 1700 and 1875" (Westendorp and Kirkwood 1998, p. 745).

The sophisticated statistical methods which these authors bring to bear on the data uncover an inverse relationship between fertility and longevity. But it is also true that a less sophisticated analysis points to the opposite conclusion.

Alternative analysis with linear regression

Naïve linear regression of age at death vs. number of children produces a correlation coefficient ($r=-0.008\pm 0.018$) that is negative but not significant. The clearest signals in the data are that over the centuries, longevity increases ($r=0.247\pm0.018$) while fertility decreases ($r=-0.126\pm0.018$). An approach based on analysis of variance suggests a two-variable regression, age at death vs number of children and year of birth, to segregate this effect. When this is done, the result is a marginally significant *positive* relationship between number of children and age at death ($r=0.025\pm 0.019$). Westendorp (Personal communication, 2001, J. Mitteldorf) has suggested that a more sensitive correction for social trends in fertility may be made if from the number of children each woman has borne is subtracted the average number of children borne by a sliding group of her contemporaries. When this is done, the positive correlation becomes significant ($r=0.060\pm0.018$).

(The DS predictions ought to apply equally well to women and men. The corresponding analysis for men in the database again yields a significant positive correlation between fertility and longevity, $r=0.021\pm0.011$.)

Figure 1 displays age at death for women, averaged by the number of their children. There is no clear downward trend in the chart, even without correction for year of birth (data for the chart and regression calculations derive from a later edition of the same source (Bloore 2000) used by Westendorp and Kirkwood, a CD ROM listing family records of

British Peerage. Included are the 2,919 women born before 1906 for whom exact birth dates and death dates are available and who survived to their 55th birthdays).

Why the disparity in conclusions?

One of Westendorp's conclusions is that a high proportion of the longest-lived women in his database were childless. This anomaly was clarified by Gavrilova and Gavrilov (2005), who conclude that the result derives from underreporting of fertility. They rechecked birth records for women reported childless in the Bloore database, and reanalyzing a larger validated database found no excess of infertile women among those with the longest life spans (Gavrilova et al. 2004).

Westendorp's only other significant results arise when the sample is limited to women born before 1700. This subsample is relatively small, numbering only 241 out of the full 2,919 in the database that meet other inclusion criteria. The group in which they claimed to detect a measurable deficit in fertility comprises the longest-lived subset of those early women, who attained the age of 80 or more. There were 26 such women, and their mean number of progeny was 2.65 ± 0.76, compared with 3.26 ± 0.28 for women who lived 55 to 79 years (uncertainties here are computed as standard deviation of the mean).

But the small sample size does not fully explain the disparity between the present results and Westendorp's: a linear regression limited to women born before 1700 produces an insignificant result ($r=-0.14\pm0.46$), and a bar chart limited to these women shows no clear trend (Fig. 2).

A more important explanation for the disparity in results is that we use linear regression while Westendorp and Kirkwood use a less familiar tool, the Poisson regression. Poisson regression can be used to measure an association between any independent variable that is continuous and a dependent variable that is Poisson distributed. The Poisson character of the dependent variable is an essential precondition; when the dependent variable is not Poisson distributed, the method can produce anomalous results. A Poisson distribution is controlled by a single parameter, and the core of Poisson regression analysis is to optimize over that parameter in such a way that likelihood of the observed distribution is a maximum.

Fig. 1 Age at death for women, averaged by the number of their children. There is no clear downward trend in the chart, even without correction for year of birth. The *line* indicates sample size, referred to the scale at the *right*

**Avg age at death vs # children
and sample size referred to scale at right
all women born before 1906**
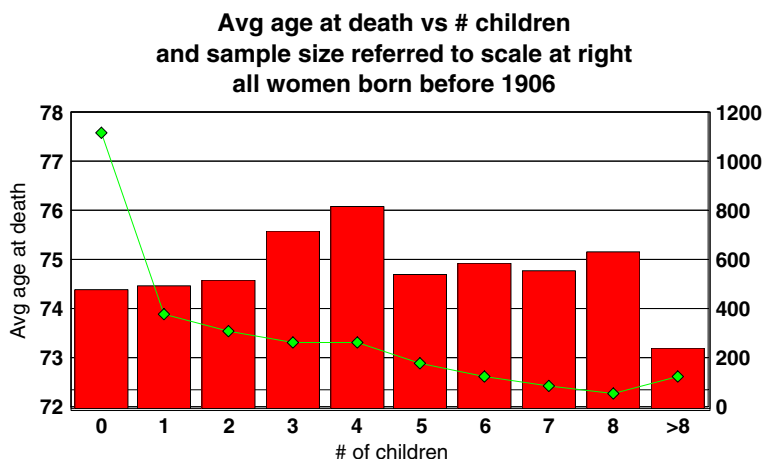
For the Westendorp analysis, the calculation would proceed thus: Assume that the probability that a given woman will bear n children derives from a Poisson probability $P(n,x)$, where the Poisson parameter x is a linear function of her age at death, $x = A \times age + B$. Try different values for A and B and calculate a product of probabilities for each woman that she would have borne the number of children that she did in fact bear. Home in on those values of A and B that maximize the probability product. Westendorp's result is that A is barely negative for women in their sample born before 1875 and is significantly negative when the sample is limited to women born before 1700. They report these results after allowance was made for trends over time in both fertility and longevity.

It is possible that a major reason for the disparity between Westendorp's Poisson result and the linear regression analysis is that number of children born to a woman is not a Poisson variable. Figure 3 shows the actual distribution of children for each woman in the sample, with a mean-adjusted Poisson curve overlaid, where $x = 2.36$ was the actual mean number of children per woman in the sample. Compared to the Poisson curve, there is an excess for $n = 0$ births and a deficiency for $n = 1$ through 4.

More important to the Poisson regression calculation is the excess of women with large numbers of children. The Poisson probability becomes exponentially small for high n, but there are substantial numbers of women in the sample with more than 10 children. For example, in the sample of 2,919, the most prolific woman bore 18 children. The Poisson probability that there should be such a woman is 10^{-7}. There are nine women with 15 or more children, while the Poisson function predicts 0.0001 such women.

Fig. 2 Same as Fig. 1, limited to women born before 1700. It is in these women that Westendorp and Kirkwood claim to discern a downward trend at the *right side* of the chart. Note the small sample size, indicated in the scale at the *right*
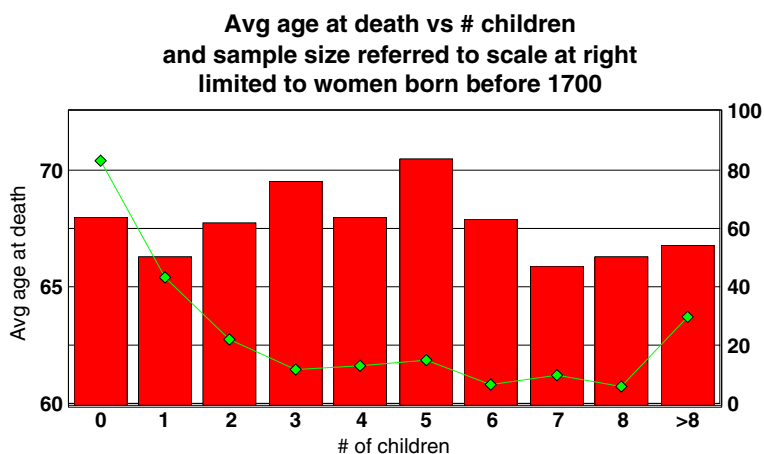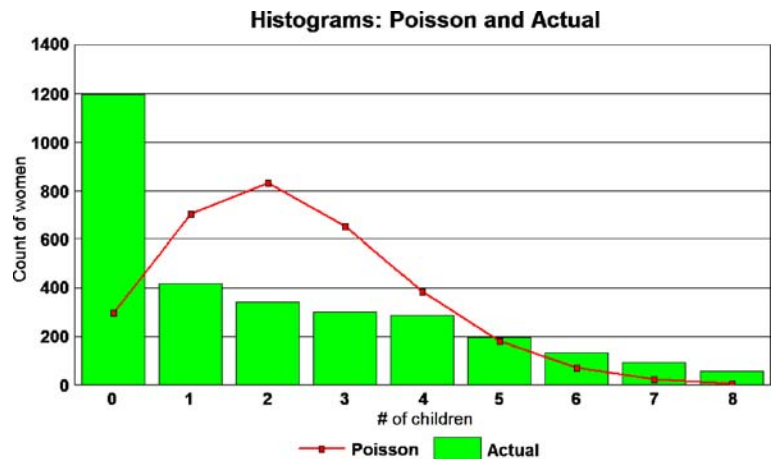
**Avg age at death vs # children
and sample size referred to scale at right
limited to women born before 1700**

Fig. 3 For Poisson regression to be a valid trend indicator, the underlying data must be Poisson distributed. The figure compares the distribution of women in the full sample by number of children with a Poisson distribution of the same mean

These observations suggest a reason for the difference between the Poisson and linear results. The Poisson probability that is maximized in Westendorp's regression calculation is inordinately influenced by the women at the very end of the distribution. Even though their numbers are small, the contribution that they make to the aggregate probability is dominant. If these few women should happen to have short lives, that would be sufficient to tilt the Poisson regression result to a negative coefficient. The average age at death of the nine women with the most children was 67.6, compared to 74.7 for the sample as a whole, a disparity that is well explained by the average birth year of the nine: 1677, compared to 1830 for the whole sample. If these nine women are removed from the sample of 2,919, the Poisson regression no longer generates a negative coefficient. Five of the nine women with the most children were born before 1700. The conclusion of Westendorp and Kirkwood depends critically on this handful of women; if they are removed from the analysis, Poisson regression on the remaining subjects produces an insignificant result, even for the pre-1700 dataset.

Laboratory evidence

These results should be viewed in the context of a growing body of animal data, which also fails to support the theoretical notion of a cost of reproduction. In his encyclopedic review of the literature, Finch (1990) reported that there was considerable animal evidence for an immediate mortality cost of reproduction but little evidence that reproduction affects the rate of aging. Stearns (1992) has appended a survey of animal tradeoff studies to his text on life histories, cataloging evidence for and against the existence of a tradeoff. Genetic experiments with nematodes suggests that genes affecting the rate of aging are not necessarily linked to fertility (Partridge and Gems 2002); while work with fruitflies has been interpreted as supportive of tradeoffs (Partridge and Gems 2002), though the data on their face demonstrate that breeding for longevity actually enhances fertility (Mitteldorf 2004). Reznick et al. (2000) review evidence that longevity and fertility in water fleas appear to be independently variable. Most recently, Ricklefs and Cadena (2007) conducted a large-scale study of 18 mammalian species and 12 avian species in captivity, and found no correlation between fertility and longevity.

Discussion

The pleiotropic theories posit that tradeoffs are the root cause of aging. In the original formulation of Williams (1957) of Antagonistic Pleiotropy, it is the potential for reproduction that is enhanced by pleiotropic alleles, so that the link between actual fertility and longevity may be less absolute. Still, we should expect to find a negative correlation in any large unbiased sample.

But the DS theory in particular predicts unequivocally that reproduction itself should have a negative impact on life span, regardless of the origin of variation in fertility. This is a central prediction that may be taken as a fair test of the theory's core.

Neither animal experiments nor human demographic data have corroborated the theoretical prediction of a strong cost of fertility.

References

Beeton M, Yule GU, Pearson K (1900) On the correlation between duration of life and the number of offspring. Proc Royal Soc B 65:290–305

Bloore J (2000) Peerage CD-ROM, 13th edn. S&N Genealogical Supplies, Salisbury

Davey Smith G, Frankel S et al (1997) Sex and death: are they related? Findings from the caerphilly cohort study. BMJ 315(7123):1641–1644

Finch CE (1990) Longevity, senescence and the genome. University of Chicago Press, Chicago

Gavrilova NS, Gavrilov LA (2005) Human longevity and reproduction: an evolutionary perspective. In: Voland E, Chasiotis A, Schiefenhövel W (eds) Grandmotherhood. New Brunswick, Rutgers University Press, pp 59–80

Gavrilova NS, Gavrilov LA et al (2004) Does exceptional human longevity come with a high cost of infertility? Testing the evolutionary theories of aging. Ann N Y Acad Sci 1019:513–517

Grundy E, Kravdal O (2008) Reproductive history and mortality in late middle age among Norwegian men and women. Am J Epidemiol 167(3):271–279

Kaplan SD (1988) Retrospective cohort mortality study of Roman Catholic priests. Prev Med 17(3):335–343

Kirkwood T (1977) Evolution of aging. Nature 270:301–304

Korpelainen H (2000) Fitness, reproduction and longevity among European aristocratic and rural Finnish families in the 1700s and 1800s. Proc Biol Sci 267(1454):1765–1770

Le Bourg E (2001) A mini-review of the evolutionary theories of aging. Is it the time to accept them? Dem Res 4:1–28

Le Bourg E, Thon B et al (1993) Reproductive life of French-Canadians in the 17–18th centuries: a search for a trade-off between early fecundity and longevity. Exp Gerontol 28(3):217–232

Lycett JE, Dunbar RI et al (2000) Longevity and the costs of reproduction in a historical human population. Proc Biol Sci 267(1438):31–35

McArdle PF, Pollin TI et al (2006) Does having children extend life span? A genealogical study of parity and longevity in the Amish. J Gerontol A Biol Sci Med Sci 61(2):190–195

Mitteldorf J (2004) Aging selected for its own sake. Evol Ecol Res 6:1–17

Muller HG, Chiou JM et al (2002) Fertility and life span: late children enhance female longevity. J Gerontol A Biol Sci Med Sci 57(5):B202–B206

Palmore EB (1982) Predictors of the longevity difference: a 25-year follow-up. Gerontologist 22(6):513–518

Partridge L, Gems D (2002) Mechanisms of ageing: public or private? Nat Rev Genet 3(3):165–175

Perls TT, Alpert L et al (1997) Middle-aged mothers live longer. Nature 389(6647):133

Reznick D, Nunney L et al (2000) Big houses, big cars, superfleas and the costs of reproduction. Trends Ecol Evol 15(10):421–425

Ricklefs RE, Cadena CD (2007) Lifespan is unrelated to investment in reproduction in populations of mammals and birds in captivity. Ecol Lett 10(10):867–872

Stearns SC (1992) The evolution of life histories. Oxford University Press, Oxford

Westendorp RG, Kirkwood TB (1998) Human longevity at the cost of reproductive success. Nature 396(6713):743–746

Williams G (1957) Pleiotropy, natural selection, and the evolution of senescence. Evolution 11:398–411