CrossMark

ORIGINAL ARTICLE

The death spiral: predicting death in *Drosophila* cohorts

Laurence D. Mueller · Parvin Shahrestani ·
Casandra L. Rauser · Michael R. Rose

Abstract *Drosophila* research has identified a new
feature of aging that has been called the death spiral.
The death spiral is a period prior to death during which
there is a decline in life-history characters, such as
fecundity, as well as physiological characters. First,
we review the data from the *Drosophila* and medfly
literature that suggest the existence of death spirals.
Second, we re-analyze five cases with such data from
four laboratories using a generalized statistical frame-
work, a re-analysis that strengthens the case for the
salience of the death spiral phenomenon. Third, we
raise the issue whether death spirals need to be taken
into account in the analysis of functional characters
over age, in aging research with model species as well
as human data.

L. D. Mueller (✉) · P. Shahrestani ·
C. L. Rauser · M. R. Rose
Department of Ecology and Evolutionary Biology,
University of California, Irvine, CA 92697, USA
e-mail: ldmuelle@uci.edu

Present Address:
P. Shahrestani
Department of Biological Science, California State University
Fullerton, 800 North State College Blvd., Fullerton,
CA 92831-3599, USA

Present Address:
C. L. Rauser
UCLA Office of Vice Chancellor for Research-UCLA
Grand Challenges, BOX 951405, 2248 Murphy Hall,
Los Angeles, CA 90095-1405, USA

Introduction

Research with humans, *Drosophila,* medflies, and
other model organisms has revealed three distinct
demographic phases (Greenwood and Irwin 1939;
Beard 1959; Carey et al. 1992; Curtsinger et al. 1992;
Vaupel et al. 1998; Mueller et al. 2003; Rose et al.
2006; Mueller et al. 2011). The first phase, which
occurs prior to the start of reproduction, is the
developmental period. Under protected conditions, at
least after the perinatal period, this phase character-
istically features low levels of mortality with no
consistent trend relative to age.

The second phase, normally called "aging," fol-
lows the onset of reproduction. During this phase, age-
specific mortality rates almost always rise, even under
protected conditions (Comfort 1979), with the singular
exception of fissile species (e.g. Bell 1984; Martinez
1998). For fissile species it could be said that there is
effectively no part of life history that follows the first
act of reproduction.

The third demographic phase, which has chiefly
been of interest since 1992 (e.g. Carey et al. 1992;
Curtsinger et al. 1992; Rose et al. 2002; Mueller et al.
2011), has been called "late life" (e.g. Rose et al.
2005; Rauser et al. 2006a). At these advanced adult

ages, the age-specific mortality rates of iteroparous species roughly stabilize, at least in human and model-species cohorts that are sufficiently large (Greenwood and Irwin 1939; Carey et al. 1992; Curtsinger et al. 1992; Vaupel et al. 1998; Rose et al. 2002; Rauser et al. 2006b).

In this paper we present a general case for the existence of a fourth feature of life cycles, which we call the "death spiral" (Rauser et al. 2005; Mueller et al. 2007, 2011). The death spiral is detectable across a wide range of adult ages, as we will show. An obvious interpretation is that it features a general and abrupt decline in physiological health prior to death. But more importantly for biogerontological analysis, death spirals create significant heterogeneity within cohorts, heterogeneity that raises problems for the analysis of age-dependent functional characters quite broadly.

Evidence for the existence of the death spiral from Mediterranean fruitflies

Papadopoulos et al. (2002) noted that 97 % of male medflies (*Ceratitis capitata*) began showing a temporary upside-down orientation, or supine behavior, about 16 days before their death. Since these medflies have an average age at death of about 62 days, the 16-day onset of medfly supine behavior is roughly equivalent to the period of time that *Drosophila* exhibit death spiral phenomena (Mueller et al. 2007 and this paper).

Müller et al. (2001) studied lifetime fecundity in 531 medflies. They noted that, after reaching a peak in early life, fecundity declined in an approximately exponential fashion. The rate of decline in fecundity was measured by a parameter, β_1, which showed wide variation among individual females. In particular, they noted that females that die early in life showed a rapid decline in fecundity with age (large β_1), while females that lived longer showed a slower decline in fecundity (small β_1). For instance, inspection of Fig. 1 in Müller et al. (2001) shows that a female that died early in life (28 days) was predicted to lay only 18 eggs at age 25 days, while the average fecundity in the whole population at this age was 27 eggs.

The Müller et al. data can be interpreted as capturing death-spiral phenomena in female fecundity. Females that die at young ages have steeper declines in age-specific fecundity, because their trajectories are based on fecundity observations that are part of the death spiral and thus decline faster than the fecundity of similarly aged females that are not about to die. Müller et al. suggest that death is a consequence of females rapidly diminishing their reproductive reserves and that these reserves vary among individuals.

Evidence for the existence of the death spiral from *Drosophila*

Studies of individual female fecundity

Over the last decade, multiple *Drosophila* laboratories have independently discovered patterns of aging that suggest the existence of death spirals among laboratory cohorts handled as adults with sufficient care to minimize the possibility of artifactual effects, such as infectious disease. In each of the relevant studies, daily counts of female fecundity were made from very early in adult life until death.

In the course of analyzing adult female *Drosophila melanogaster* handled individually, Rauser et al. (2005) found that individual female fecundity steeply declines just prior to death. Further quantitative analysis of the Rauser data led us to develop a formal analysis of the death spiral (Mueller et al. 2007, 2011). This analysis incorporated a two-phase adult female life-history, with both aging and late life, as well as a distinctive death spiral phase in which fecundity declines linearly, but at a more rapid pace than the decline that characterizes normal aging.

Rogina et al. (2007) inferred the existence of a comparable death spiral phenomenon for *Drosophila* female fecundity, noting that females that were about to die the next day always laid zero eggs and that independently of their mating schedule showed fecundity declines for a week or more before death. The egg-laying of females on the day before that zero-fecundity day averaged about 0.2 eggs. Their study used 386 females from a mutant balancer stock of *D. melanogaster*.

Curtsinger (2015) proposed that the age at which a female first lays zero eggs is a significant indicator of impending death. In his analysis of several previously published databases, specifically those of Le Bourg et al. (1988), Rauser et al. (2005), Klepsatel et al.

(2013) and Khazaeli and Curtsinger (2014), he shows that roughly 10 days before the first zero egg count day a standardized fecundity measure starts declining (Curtsinger 2015; Fig. 3b), which corresponds roughly to the start of the death spiral that had been previously inferred by Rauser et al.

Death spirals in male virility data

Shahrestani et al. (2012a) studied male fertility, often called virility, by counting the number of females, of eight total, that a male could fertilize in 24 h. They found that the virility of males that were within 7 days of death was significantly lower than that of similarly aged males that were not about to die.

Materials and methods

Lifetime female fecundity

In the first re-analysis, we relied on lifetime fecundity records collected from the following four sets of cohorts: (Le Bourg et al. 1988; Rauser et al. 2005; Klepsatel et al. 2013; Khazaeli and Curtsinger 2014). We first review some important details of the biological Materials and methods used in the construction of each of these databases.

Le Bourg et al. (1988) collected lifetime fecundity records on 322 females. These samples came from three populations, two were selected for activity level and the third was a control population. Assays were conducted over seven generations of selection. These populations were started from a standard Oregon R stock and at each generation only ten pairs of flies were used to reconstitute the next generation. Over the course of this experiment, heterozygosity was reduced by about 17 %, and in all likelihood the original Oregon R stock was inbred to some degree. Vials consisted of a male–female pair and males were not replaced upon death. The flies were given standard food with live yeast.

Khazaeli and Curtsinger (2014) used two inbred lines, R17 and S9, to derive a total of 335 females. These inbred lines were created by 28 generations of full sib mating using two populations created by Luckinbill and Claire (1985). The R17 inbred line was derived from Luckinbill and Claire's LA line which was selected for late life fitness, while the S9 inbred line was derived from Luckinbill and Claire's LD1 control line. Females were assayed in vials with two males that were replaced as needed. Fresh food was supplied daily, although there did not appear to be any yeast supplementation.

Klepsatel et al. (2013) collected lifetime fecundity records on 488 females derived from three populations collected from (i) Austria, (ii) South Africa, and (iii) Zambia. These populations were maintained in the laboratory for only 2–4 generations prior to these assays. The Austrian population was created from a sample of 200 flies, the South African population from 140 flies derived from 7 isofemale lines, and the Zambian population from 600 flies derived from 30 isofemale lines. An isofemale line would typically undergo at least one generation of full sib mating. Females were kept in vials with about 10 mg of live yeast supplement and 2 males. Dead males were replaced.

Rauser et al. (2005) used adult flies that were aged 12 days from egg (not 12 days as adults as reported by Curtsinger 2015), which made the adults 1–2 days old at the start of the assays. Each female was housed in a vial with 2 males and these males were replaced as they died. Vials were supplemented with 5 mg live yeast. A single, large, lab-adapted population (called "CO_1" created in the study of Rose et al. 1992) was used to derive three independent cohorts, yielding a total of sample size of 2828 females.

Lifetime male virility

We analyzed lifetime male virility from the data of Shahrestani et al. (2012a). Virility observations consist of counts of the number of females that a male fertilizes in 8 h. Like with the fecundity data, Shahrestani et al. tested large samples (712) of males once a week until their death.

Scaled age-specific fitness

The concept of the death spiral is a death process. Accordingly, we sought a means of examining this process independent of an individual's precise age. We chose to recast an individual's age as a series of target ages. We first describe the method used for female fecundity. At each target age we could classify a female as either alive or dead by the end of the target age. This

would then create two groups of females and each female would have an associated history consisting of her record of eggs laid $0, 2,\ldots, m$ days prior to the target age, e.g. days before death. To standardize an individual's fecundity phenotype, we scaled and centered the entire population's fecundity measurements at each age. Thus, if at age-t, the average fecundity of all live females and females that died at age-t were, f_t, and the standard deviation, σ_t, then the fecundity of female-j at time-t, f_{tj}, would be transformed to $\tilde{f}_{tj} = \frac{f_{tj}-f_t}{\sigma_t}$. If female-$j$ is found dead on day-t there is a 24 h window on her exact time of death. Thus, the f_{tj} values may reflect the eggs produced in a 24 h period (if the female died just before the census), or perhaps the eggs laid in just a few minutes of time (if she died immediately after being transferred to the fresh vial).

If at time-t there are n_t, females at that age, either alive or those that died that day, we generate predictor vectors for female-j ($j = 1,\ldots, n_t$) that look like, $\mathbf{f}_j = (\tilde{f}_{m,j}, \tilde{f}_{m-1,j}, \ldots, \tilde{f}_{0,j})$, where we have now transformed the time scale so $t = 0$, $t-1 = 1$, and so on. Thus, $\tilde{f}_{k,j}$, is the transformed fecundity for female-j, k days before the target age. This process is repeated for the next age, $t + 1$, and $n_{t + 1}$ new records are created (Fig. 1). Obviously, females that did not die by age $t + 1$ will have nearly the same fecundity records entered into the database at both target ages t and $t + 1$. When creating these records, we omitted

Four-day window analysis of data from a single female

					Target Ages		
				↓	↓	↓	↓
Adult age	21	22	23	24	25	26	27
Scaled fecundity	0.9	1.3	0.4	0.1	-1.2	-1.7	NA
Female status	live	live	live	live	live	dead	

Window A. live female status, (0.9, 1.3, 0.4, 0.1)
Window B. live female status, (1.3, 0.4, 0.1, -1.2)
Window C. dead female status, (0.4, 0.1, -1.2, -1.7)

Fig. 1 An example of scaled fecundity records from a single female. The scaled fecundity values for one female at ages 2–26 days are shown using a 4 day window. The female is found dead at the end of day 26. The records generated at target days 24, 25 and 26 from this female are shown. The record for target day 24 is labelled *A*, for day 25, *B* and for day 26, *C*. Since the female dies on day 26 no record for this female is generated at day 27. In this example a single female generated two records classified as "live" and one classified as "dead". In general each female will have many live records and only a single dead record
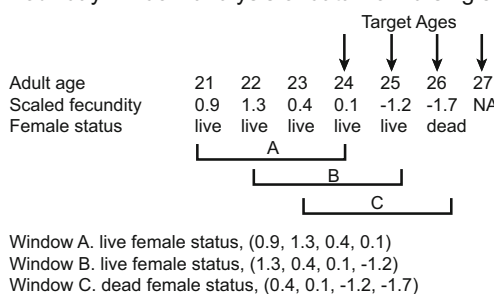
females that did not have complete fecundity records for the entire window of $m + 1$ days. In addition, females that died at ages $<m$ were not included in these records. These records were used to do regressions of scaled fecundity on days before the target age for the dead females and the live females. Since the database of live females included the same females multiple times, observations are correlated. Therefore, the analyses of the regression results must use methods that can take into account these correlations.

We did bootstrap resampling of our scaled fecundity dataset to determine the variances in scaled fecundity at various ages before the target age and to determine if the slope of the regression lines for the dead and live females differed. The total size, N, of the Rauser et al. dataset was large: 43,093 with a 15-day window ($m = 14$). At each bootstrap sample, we sampled with replacement N female fecundity records. From this we took a subsample of 1000 records of live females and 1000 records of dead females. We then computed the mean scaled-fecundity of each group and the regression of scaled fecundity versus days before target age. We did 100 bootstrap samples and from these computed empirical 95 % confidence intervals on the mean scaled-fecundity and confidence intervals on the slopes of each line.

Some females will stop laying eggs sporadically or completely prior to death. Curtsinger (2015) has recently argued that the appearance of days with zero egg counts is an important indicator of the end of life, a phase he has called the retired stage. We followed up on his suggestion by examining how useful zero egg counts are for predicting death. To do this we took the records of lifetime female fecundity and changed each daily egg count to 0 if no eggs were laid and 1 if one or more eggs were laid. We then used the target age concept described previously to generate records for live and dead females.

The virility data at each age was scaled and centered in the same way we transformed female fecundity data. We also did bootstrap re-sampling to estimate the variance of the scaled virility values of males that were alive and dead on the target day.

Predicting death at the target age

Can we use scaled fecundity or zero egg counts at say up to m-days before the target age to predict which female will be dead at the target age? This is not as

simple as one might think. Recall that a female that dies one day after the target age will be included in the "live" group and her fecundity record will be the same as when she actually dies, except for the inclusion of new information at day 1 and the removal of the day m observation.

We used a statistical learning technique, boosted classification trees, to predict whether a female is dead or alive on the target age based on her fecundity record in the previous m days. Classification trees provide predictions on an individual's membership in a binary classification based on successive partitioning of the predictor space (Hastie et al. 2009, Chap. 9). Boosting is a technique for improving the prediction of methods such as classification trees, which employ the sequential building of trees in which each step creates additional tree branches that are constructed to improve the prediction of the residuals from the previous tree (Hastie et al. 2009, Chap. 9).

In order to avoid developing a classification tree that focused on the correct prediction of live females due simply to the much larger number of records for live flies, we created a data set with an equal number of records of live and dead flies from the Rauser et al. data. From this we used 80 % of all the data (training data) to train the boosted classification tree. The remaining 20 % of the observations (testing data) were used to predict whether each female was alive or dead at the target age. We determined how many days prior to the target age provided the best predictions based on 10-fold cross validation of the training data set. In addition, the same cross-validation statistics were used to determine (i) the number of trees that resulted in a minimum error, (ii) the best depth of the individual trees, and (iii) the optimal shrinkage parameter (for a discussion of these issues see Gareth et al. 2013, Chap. 8).

Predictions with the testing data gave rise to two sorts of errors. Females that were dead at the target age were incorrectly predicted to be alive (dead error rate), and females that were alive at the target age were incorrectly predicted to be dead (alive error rate). If we simply guessed female status the error rates would be 50 % since we had equal numbers of live and dead females in the testing data.

Destructive physiological measurements

If fitness components like fecundity and virility decline during the death-spiral, then we expect other physiological functions to change prior to death as well. One problem with testing this idea is that in *Drosophila* many physiological tests are destructive, and therefore we cannot use death as an indication of which flies are in the death-spiral. We have proposed as a solution to this problem to use female fecundity for several days prior to a destructive physiological assay (Mueller et al. 2009; Shahrestani et al. 2012b). One difficulty with any potential test by these techniques is that the physiological character may be correlated with fecundity, and thus by using fecundity to identify females in the death-spiral we may get a sample that looks different simply because of these correlations.

Shahrestani et al. (2012b) developed a procedure that utilized principal components of physiological characters, such as desiccation resistance and time-in-motion, rather than the raw measurements. Thus, for each female the centered and scaled desiccation and time-in-motion observations were transformed to a vector, $z_i = (z_{1i}, z_{2i})$, $i = 1,..,n$, of the two principal components. Shahrestani et al. partitioned the n principal component vectors into a "death-spiral" group, with principal components $\mathbf{Z}_s$, and a "non death-spiral" group, with principal components $\mathbf{Z}_{ns}$, according to each female's fecundity. Using the mean, $\hat{\mu}$, and covariance, $\hat{\Sigma}$, of $\mathbf{Z}_{ns}$ the Mahalanobis distance for every female was estimated as, $d_i^2 = (z_i - \hat{\mu})^T \hat{\Sigma}^{-1}(z_i - \hat{\mu})$. The test statistic is then a simple t-statistic resulting from a comparison of the mean Mahalanobis distance among spiral females versus the mean among non-spiral females. This statistic has a nominal 5 % type-I error rate even when there are correlations between fecundity and these characters, but no difference in the mean value of these traits in the two groups (Shahrestani et al. 2012b).

Results
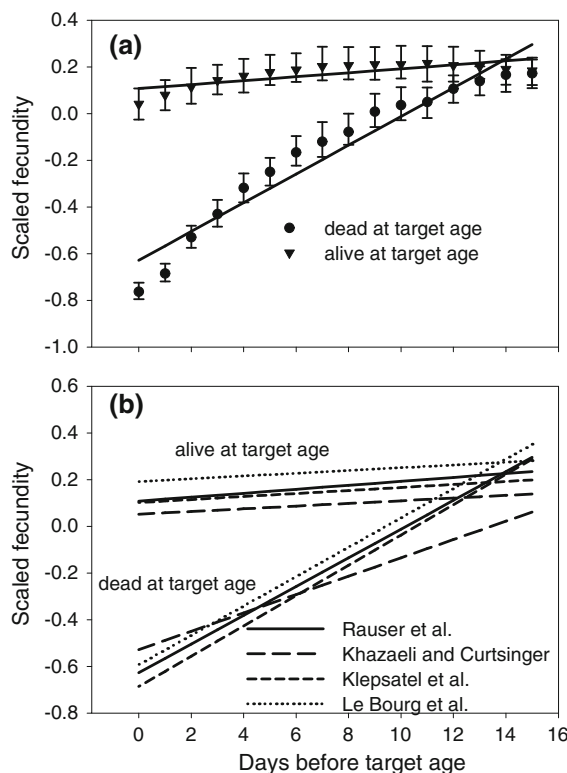
Lifetime female fecundity

For the Rauser et al. data set, we see that the difference between the scaled fecundity of the dead and alive females becomes greater the closer the comparison is to the target age (Fig. 2a). The slopes of the regression are significantly different. Thus, we get an indication from these data that female fecundity starts declining

well before death, perhaps as early as 2 weeks before. This is a strong indication that flies that are dying from intrinsic causes, not accidents or catastrophes, start to show physiological manifestations of this well before death. As we will see later, female fecundity is not the only trait that is adversely affected by the death process.

How robust are these results? We address this by creating scaled fecundity data for the three other published *D. melanogaster* data sets described previously. Each of these shows remarkable consistency with the Rauser et al. results (Fig. 2b), suggesting that the death spiral is common to this species, at least for female fecundity.

Although we have suggested that the death spiral is not a direct product of aging, its characteristics may be
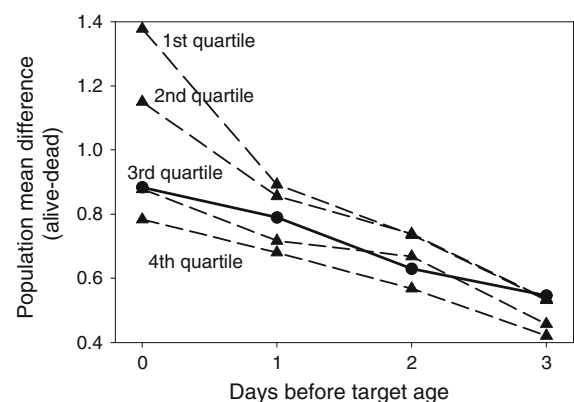
a function of age. To study this, we divided the females from Rauser et al. into quartiles according to their ages at death. Thus, the first quartile consisted of scaled fecundity records for females at target ages less than 24 days, then 24–30 days, 31–38 days, and greater than 38 days for the second, third and fourth quartile, respectively. When we examine the difference between the mean scaled fecundity for 4 days prior to the target day (Fig. 3), we see that age at death has a large effect on the difference between the live and dead groups at the target age, but that the differences between quartiles diminishes rapidly as we move further from the target age. Females that die at a young age will be more impacted by the death spiral than females that die at older ages, especially at ages very close to death. So while age at death has some impact on the magnitude of the death spiral effect, the number of days before death is a more substantial indicator of the impact of the death spiral effects.

Analysis of zero egg counts

Using the results from all four laboratories, we see that there is a distinct difference for the fraction of females laying zero eggs when we compare the females alive at the target age with those dead at the target age (Fig. 4). As mentioned previously, since some females may have only short periods to lay eggs on the day they died, the day zero observations are of limited utility. Just before death, all populations show a very high

Fig. 2 a Rauser et al. scaled fecundity for females dead at target age (*circles*) and alive at the target age (*triangles*) for up to 16 days before the target age. The *bars* are 95 % confidence intervals calculated from 100 bootstrap samples. The slope of the *line* through the live female data is 0.0085, with a 95 % confidence interval of (0.004, 0.013) and the dead female slope is 0.062, with a 95 % confidence interval of (0.058, 0.066). **b** Regression lines for the Rauser et al. data in (**a**) and three other data sets as indicated

Fig. 3 The differences between scaled fecundity in the live and dead females as a function of age-at-death (*dashed lines*). The first quartile of data represents the youngest females while the fourth quartile represents the oldest females. The *solid line* is for the entire Rauser et al. data set

fraction of females laying no eggs, although it is not 100 %, and the value varies substantially between lab populations. In the Rauser et al. and Khazaeli and Curtsinger cohorts, about 60 % of the females are not laying eggs the day before death, while in the Le Bourg et al. and Klepsatel cohorts, the number is closer to 80 % and as reported previously is 100 % for the flies studied by Rogina et al. (2007). As with the scaled fecundities, the fraction of females with zero egg counts for females dead on the target day declines as we look further from the day the fly dies. It is similar to the values for live flies by about day 14 before death. In this sense, the zero egg count days shows a death spiral decline that has a similar dynamic to the scaled fecundity decline.

We used the large data set of Rauser et al. to examine if the fraction of females laying zero eggs differs depending on when a female died. We divided the Rauser et al. data into quartiles, as we did with the scaled fecundity. We see that the fraction of females with zero eggs is generally higher for females that die at older ages (Fig. 5). It also appears that as we examine days further removed from the day of death, fewer females are laying zero eggs and this is most pronounced with young females.

The best predictions utilized only 4 days of scaled fecundity records prior to the target day. For the Rauser et al. test data, the best predictions were for the dead females (Fig. 6a) and the average error rate was about 23 %. Using the boosted classification tree trained on the Rauser et al. scaled fecundity data, we

also predicted the status of females in the three other data bases. While the predictions are slightly better for the Rauser et al. data, there were low prediction errors in all the other databases (Fig. 6a), and the predictions were certainly much better than simply guessing.

We next repeated the construction of boosted classification trees using the zero egg count data. Even though the classification tree was trained on the Rauser et al. data, the Le Bourg et al. and Klepsatel et al. data had a lower error rate for dead females than the Rauser et al. data (Fig. 6b). We can make sense out of these results if we examine the relative importance of the predictor variables. Each dependent variable can be assigned a relative weight in a classification tree based on the frequency it is used to bifurcate each branch of the tree. The day before the target age provides the most important predictor data when using either scaled fecundity or zero egg counts, although it is more so with zero egg counts (Fig. 7). The Klepsatel et al. and Le Bourg et al. data have a higher frequency of zero egg counts the day before death (Fig. 4) than the other two databases, accounting for the more accurate prediction of which females will die. Conversely, the Rauser et al. data show the lowest frequency of zero egg counts the day before the target age among live females (Fig. 4) and hence the most accurate prediction of which females are alive (Fig. 6b).

Are zero egg counts better at prediction than scaled fecundity? We can use the dead female error rates to
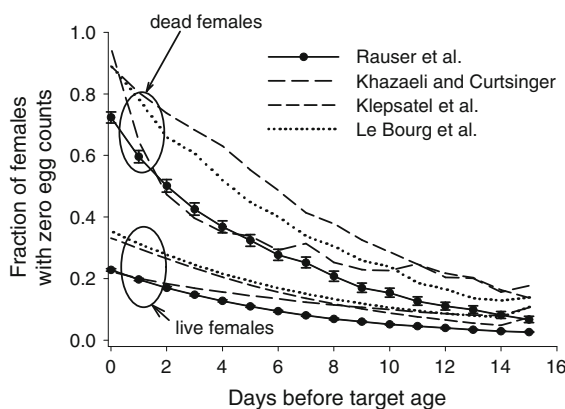
Fig. 4 The fraction of females laying zero eggs as a function of time before target ages in four different *Drosophila* databases. The *bars* on the Rauser et al. *lines* are 95 % confidence intervals computed from a standard binomial distribution
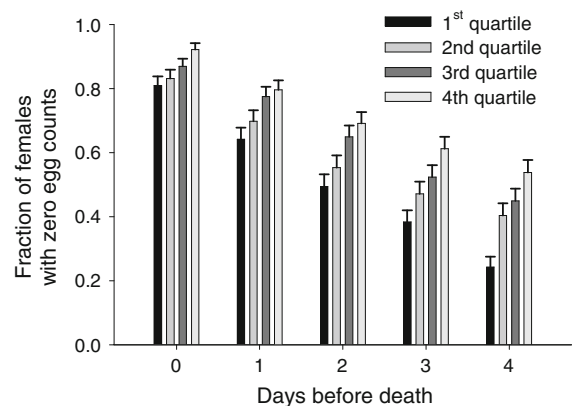
Fig. 5 The fraction of zero egg counts as a function of time before death in the Rauser et al. data. The data has been stratified by age at death. Thus, the first quartile represents the zero egg counts for females that were among the first 25 % to die and so on. The *bars* are 95 % confidence intervals
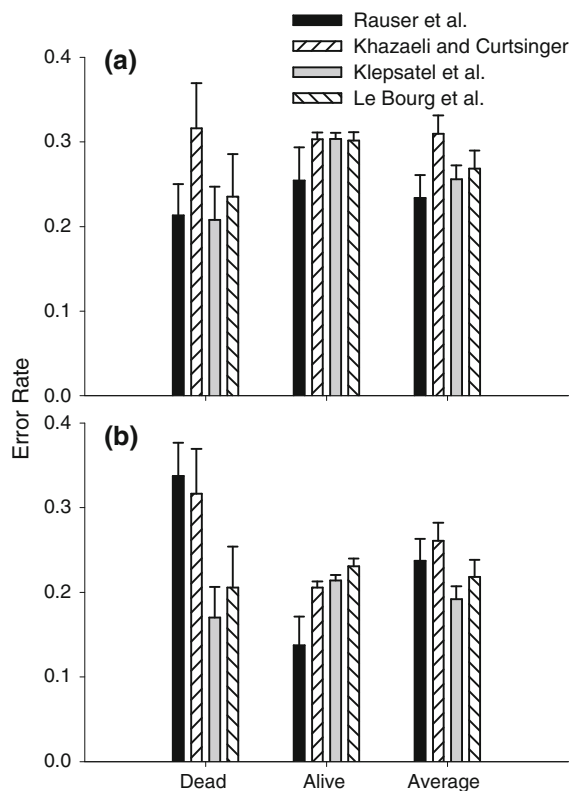
Fig. 6 Error rates from boosted classification trees in *four different databases*. The *bars* are 95 % confidence intervals. **a** Error rates for the scaled fecundity data. **b** Error rates for the zero egg count data. See the text for definitions of the error rates
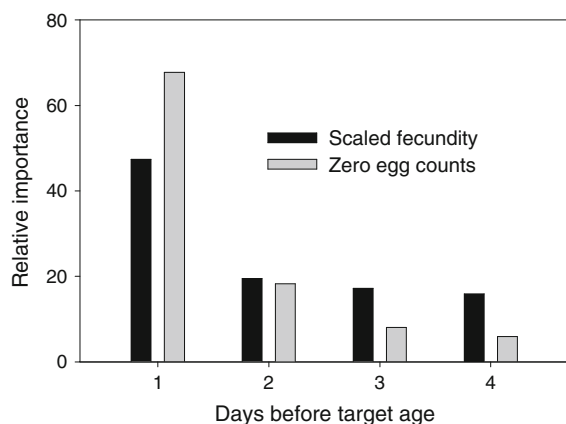
Fig. 7 The relative importance of the independent variables, 1, 2, 3, and 4 days before the target age in the construction of boosted classification trees when using scaled fecundity (*black bars*) and zero egg counts (*grey bars*). The classification tree was trained with the Rauser et al. data. The relative importance values sum to 100 and represent the percent of the bifurcations using each variable

do simple χ^2 tests of dead error rates in each database comparing the predictions based on scaled fecundity records to zero egg count records. We used only the dead error rates, since these are constructed from independent female data, while the live error rates are not. Only the Rauser et al. data show a significantly different dead error rate. The scaled dead error rate (21 %) in the Rauser et al. database was significantly lower than the zero egg count rate (34 %, $p < 10^{-5}$).

Curtsinger (2015) has suggested that the days when flies lay zero eggs are especially important for identifying females that are characterized by low levels of oviposition and reduced prospective survival. Curtsinger uses the first zero-egg day as an indicator of the start of a life phase he calls retired. Females not in the retired stage are in the working stage. We have shown that the proportion of females that have zero egg days shows a pattern similar to average fecundity (Fig. 4). That is, prior to death there is, on average, about a two-week period during which the chance of a female with a zero egg day increases. We suggest that zero egg count days are simply a different scaling of fecundity and that the patterns of zero egg days can be explained by the death spiral.

We find the retired/working dichotomy less useful than the death spiral. Firstly, the retired/worker classification only works for females. But as we have shown, males undergo a similar decline in reproductive function prior to death. The death spiral includes multiple physiological traits, not simply reproductive traits. As a classification scheme, the retired/worker scheme is highly variable. In the Rauser et al. data, 20 % of females show no zero egg counts up to and including the day before they die. Among females showing a zero egg count day before death, the median first zero egg count day occurs seven days before a female's death. Thus, for females with zero egg count days, 50 % of them will have their first zero egg count day half way into the death spiral. Lastly, for predicting female death, zero counts work as well as scaled fecundity in 3 out of 4 databases, but in the Rauser et al. database scaled fecundity is significantly better than zero egg counts. It should be noted that the testing set of data used to predict the dead error rates in the Rauser et al. database had only about 500 samples, which was not substantially higher than the sample sizes used for predicting dead error rates in the other databases. Hence, the observed significance in the Rauser et al. database cannot be due to simply a difference in power.
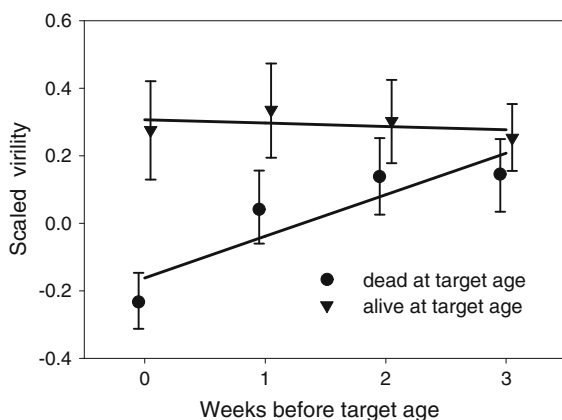
Male virility

The slope of scaled male virility vs. weeks before target age (Fig. 8) for alive males is small (−0.01) and not significantly different from zero (95 % confidence interval, (−0.067, 0.046)). On the other hand, the males that were dead within a week of the target age showed a positive slope (0.12) that was significantly different from 0 (95 % confidence interval, (0.082, 0.16)). The pattern is very similar to that of fecundity, as displayed in Fig. 2, suggesting that the death spiral results in declining male virility as early as 2 weeks before death.

Desiccation resistance and time-in-motion

The distribution of distance values in the two groups (Fig. 9) shows that there is a higher frequency of small distance values in the non-spiral group and more frequent large distance values in the spiral group. The result of these differences in the distribution is that the two group means are significantly different ($p = 0.00069$). The p value reported in Shahrestani et al. (2012b) was for a two-tailed test, although in fact the correct alternative hypothesis is that the average distance in the spiral group is greater than the average distance in the non-spiral groups, which was in fact what was observed. Both mean time-in-motion and desiccation resistance were reduced in the death spiral group, indicating a decline in physiological function.

Fig. 8 Scaled virility for males that dies within a week of the target age (*circles*) and those that were alive for at least a week after target age (*triangles*) for up to 3 weeks before the target age. The *bars* are 95 % confidence intervals calculated from 1000 bootstrap samples
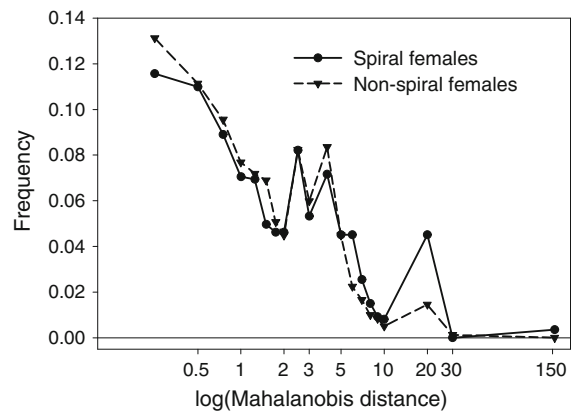
Fig. 9 The distribution of Mahalanobis distance values based on principal components of desiccation resistance and time-in-motion measurements in females identified as either in the death spiral or not in the death spiral based on their fecundity prior to assays

While the distributions (Fig. 9) look similar, the tests on the means used large samples to overcome the inevitable misclassification problem with data of this kind. Shahrestani et al. tested 3272 females for both desiccation resistance and time-in-motion.

Discussion

General lessons for the construction of demographic models

One natural consequence of the death spiral is that population-level measurements of changes in phenotypes with age must allow for possible effects produced by the death spiral. In principle, a phenotype could stay constant in an individual until they enter the death spiral, at which point there would be a consistent and severe decline. At the population level, it would appear as if the average value of this phenotype declines with age, since an increasing fraction of the population would be expected to be in the death spiral as a cohort of individuals get older. In the case of fecundity and male virility, we know that such a pattern of non-aging is not the case, because after removing the individuals in the death spiral, we still see an age dependent decline in these phenotypes (Mueller et al. 2007; Shahrestani et al. 2012a). However, other phenotypes would have to be individually tested for the impact of death spirals.

To model the dynamics of age-specific change in a phenotype like average fecundity, the number of females in the death spiral and the number that are not would need to be accounted for. Mueller et al. (2007) demonstrated how the death spiral can be incorporated into a model of age-specific fecundity. We should note that the population heterogeneity generated by the death spiral is fundamentally different from the heterogeneity in age-specific mortality that has been invoked as an explanation of mortality plateaus (Vaupel et al. 1979; Mueller et al. 2011, Chap. 8). There is no need to invoke heterogeneity in lifelong robustness to generate death spiral phenomena.

Müller et al. (2001) argue that lifespan is tied to the rate at which fecundity declines and this rate is an indicator of frailty. They reason that females showing rapid fecundity declines early in life are frailer, and thus can be expected to die at an early chronological age. The death spiral, on the other hand, suggests the death process affects many physiological traits. With this interpretation, the rapid decline in female fecundity early in life is a symptom of dying rather than a cause of death. Traits affected by the death spiral may be related to important fitness characters, like fecundity and virility, or may be traits that are relatively unconnected with fitness, like supine behavior or time-in-motion. Further research is required to sort through these alternative explanations of the death process.

A review of the literature suggests that many previously documented biomarkers of aging or death may in fact be part of the death spiral. We review some of these next.

Wax and Goodrick (1978) have reported that old mice within 1 week of death display more random wheelrunning than old mice that are not within a week of death. Recently, Belsky et al. (2015) suggested that humans in midlife that are aging more rapidly also show (i) less physically ability, (ii) cognitive decline and brain aging, (iii) self-reported worse health, and (iv) looked older. Zhang and Pincus (2016) report that blood pressure and BMI are predictive of mortality in mid-life, while blood glucose is predictive at later ages.

Rera et al. (2012) recently reported that adult *D. melanogaster* fed food laced with blue dye become blue colored sometime shortly before death. Rera et al. call these blue individuals "Smurfs." The fly's uptake of dye is related to the loss of intestinal barrier function. Smurfs also exhibit increased antimicrobial peptide expression, and impaired insulin/insulin-like growth factor signaling. It seems likely that the Smurf phenotype and its associated physiological syndromes are also part of the death spiral, although additional experiments would be needed to confirm this.

In longitudinal assays of individual *D. melanogaster* using real-time video tracking of GFP fluorescence, hsp22 and hsp70 transgenic reporters began to spike in expression $\sim$5–10 h before death (Grover et al. 2008, 2009). It has been suggested that hsp gene expression levels could possibly be used to predict remaining life span of individuals (Yang and Tower 2009).

The case of human disability

Chronic disability among the U.S. population has been declining since 1982 (Cutler 2001; Manton and Gu 2001; Crimmins 2004; Manton et al. 2006). This decline has occurred despite the fact that the elderly have been increasing in numbers and living longer. Numerous reasons have been proposed for this decline, including improved human physiology from better diets and better medical care (Fogel and Costa 1997; Schoeni et al. 2008). Manton et al. (2006) estimate that these declines in disability rates saved Medicare around $26 billion in 1999 alone. The enormous costs of assisting the disabled make the study of the process of disability an important practical problem.

It has been suggested that, prior to becoming disabled, individuals go through a series of steps called the disablement process (Verbrugge and Jette 1994). The disablement process suggests that some pathology, like a disease or an injury, may lead to an impairment that is a dysfunction of a specific body system. Examples of such body systems include the musculoskeletal and cardiovascular systems. These impairments can lead to a specific limitation in activity or movement, which becomes a disability when it limits the individual's ability to carry out activities of daily life. This description of the disablement process has found support in the bio-medical literature (Fauth et al. 2008). The onset of disability may lead to a downward spiral of new pathologies and ultimately to death (Verbrugge and Jette 1994; Morley 2008). We suggest that this process in humans is the analog of the death spiral we have documented in fruit flies.

Support for this last conjecture has been provided by Christensen et al. (2008). They monitored the physical and cognitive abilities of 2262 Danish individuals, all born in 1905. Over the course of the study, the individuals were between 92 and 100 years of age. They found that the physical and cognitive scores of a group of individuals that died within 2 years of the measurements were significantly lower than the scores of similar aged individuals who did not die.

At present we do not understand the environmental and genetic factors that may affect the onset and duration of the period of disability prior to death. Research on this problem would be difficult, expensive, and time consuming on human subjects. Appropriate animal model research would be a more effective approach to resolving such questions. In particular, lab cohorts that have been partitioned with respect to the death spiral may provide useful information with respect to the biochemical physiology which underlies the transition to the death spiral. Finally, it is not clear from the present analysis whether entry into the death spiral is irreversible or not. As this is ultimately a biotechnological question, we are hopeful that further research will yield methods of rescuing individual organisms from the death spiral.

References

Beard RE (1959) Note on some mathematical mortality models. In: Wolstenholme GEW, O'Connor M (eds) The lifespan of animals, Ciba Foundation Colloquium on Ageing. Little, Brown, Boston, pp 302–311

Bell G (1984) Evolutionary and nonevolutionary theories of senescence. Am Nat 124:600–603

Belsky DW, Caspic A, Houts R, Cohen HJ, Corcoran DL, Danese A, Harrington H, Israel S, Levine ME, Schaefer JD, Sugden K, Williams B, Yashin AI, Poulton R, Moffitt TE (2015) Quantification of biological aging in young adults. Proc Natl Acad Sci USA 112:E4104–E4110

Carey JR, Liedo P, Orozco D, Vaupel JW (1992) Slowing of mortality rates at older ages in large medfly cohorts. Science 258:457–461

Christensen K, McGue M, Petersen I, Jeune B, Vaupel JW (2008) Exceptional longevity does not result in excessive levels of disability. Proc Natl Acad Sci USA 105:13274–13279

Comfort A (1979) Ageing: the biology of senescence, 3rd edn. Churchill Livingstone, Edinburgh

Crimmins EM (2004) Trends in the health of the elderly. Ann Rev Public Health 25:79–98

Curtsinger JW (2015) The retired fly: detecting life history transition in individual *Drosophila melanogaster* females. Gerontol A Biol Sci Med Sci 70:1455–1460

Curtsinger JW, Fukui HH, Townsend DR, Vaupel JW (1992) Demography of genotypes: failure of the limited life span paradigm in *Drosophila melanogaster*. Science 258:461–463

Cutler DM (2001) The reduction in disability among the elderly. Proc Natl Acad Sci USA 98:6546–6547

Fauth EB, Zarit SH, Malmberg B (2008) Mediating relationaships within the disablement process model: a cross-sectional study of the oldest-old. Eur J Aging 5:161–179

Fogel RW, Costa DL (1997) A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs. Demography 34:49–66

Gareth J, Witten D, Hastie T, Tibshirani R (2013) An introduction to statistical learning with applications in R. Spinger, New York

Greenwood M, Irwin JO (1939) Biostatistics of senility. Hum Biol 11:1–23

Grover D, Yang J, Tavare S, Tower J (2008) Simultaneous tracking of fly movement and gene expression using GFP. BMC Biotechnol 8:93

Grover D, Yang J, Ford D, Tavare S, Tower J (2009) Simultaneous tracking of movement and gene expression in multiple *Drosophila melanogaster* flies using GFP and DsRED fluorescent reporter transgenes. BMC Res Notes 2:58

Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning, 2nd edn. Springer, New York

Khazaeli AA, Curtsinger JW (2014) Heterogeneity's ruses: how hidden variation affects population trajectories of age-dependent fecundity in *Drosophila melanogaster*. Demogr Res 30:313–332

Klepsatel P, Galikova M, De Maio N, Ricci S, Schlotterer C, Flatt T (2013) Reproductive and post-reproductive life history of wild-caught *Drosophila melanogaster* under laboratory conditions. J Evol Biol 26:1508–1520

Le Bourg E, Lints FA, Delince J, Lints CV (1988) Reproductive fitness and longevity in *Drosophila melanogaster*. Exp Gerontol 23:491–500

Luckinbill LS, Clare MJ (1985) Selection for life span in *Drosophila melanogaster*. Heredity 55:9–18

Manton KG, Gu X (2001) Changes in the prevalence of chronic disability in the United States black and nonblack population above age 65 from 1982 to 1999. Proc Natl Acad Sci USA 98:6354–6359

Manton KG, Gu X, Lamb VL (2006) Change in chronic disability from 1982 to 2004/2005 as measured by long-term

changes in function and health in the U.S. elderly population. Proc Natl Acad Sci USA 103:18374–18379

Martinez DE (1998) Mortality patterns suggest lack of senescence in Hydra. Expt Geront 33:217–225

Morley JE (2008) Diabetes and aging: epidemiologic overview. Clin Geriatr Med 24:395–405

Mueller LD, Drapeau MD, Adams CS, Hammerle CW, Doyal KM, Jazayeri AJ, Ly T, Beguwala SA, Mamidi AR, Rose MR (2003) Statistical tests of demographic heterogeneity theories. Exp Gerontol 38:373–386

Mueller LD, Rauser CL, Rose MR (2007) An evolutionary heterogeneity model of late-life fecundity in *Drosophilia*. Biogerontology 8:147–161

Mueller LD, Shahrestani P, Rauser CL (2009) Predicting death in female Drosophila. Exp Gerontol 44:766–772

Mueller LD, Rauser CL, Rose MR (2011) Does aging stop? Oxford University Press, Oxford

Müller HG, Carey JR, Wu D, Liedo P, Vaupel JW (2001) Reproductive potential predicts longevity of female Mediterranean fruitflies. Proc R Soc Lond Ser B 268:445–450

Papadopoulos NT, Varey JR, Katsoyannos BI, Muller HG, Liu X (2002) Supine behavior predicts time-to-death in male Mediterranean fruit flies. Proc R Soc Lond Ser B 269:1633–1637

Rauser CL, Abdel-Aal Y, Shieh JA, Suen CW, Mueller LD, Rose MR (2005) Lifelong heterogeneity in fecundity is insufficient to explain late-life fecundity plateaus in *Drosophila melanogaster*. Exp Gerontol 40:660–670

Rauser CL, Mueller LD, Rose MR (2006a) The evolution of late life. Aging Res Rev 5:14–32

Rauser CL, Tierney JJ, Gunion SM, Covarrubias GM, Mueller LD, Rose MR (2006b) Evolution of late-life fecundity in *Drosophila melanogaster*. J Evol Biol 19:289–301

Rera M, Clark RI, Walker DW (2012) Intestinal barrier dysfunction links metabolic and inflammatory markers of aging to death in Drosophila. Proc Natl Acad Sci USA 109:21528–21533

Rogina B, Wolverton T, Bross TG, Chen K, Muller HG, Carey JR (2007) Distinct biological epochs in the reproductive

life of female *Drosophila melanogaster*. Mech Ageing Dev 128:477–478

Rose MR, Vu LN, Park SU, Graves JL (1992) Selection for stress resistance increases longevity in *Drosophila melanogaster*. Exp Gerontol 27:241–250

Rose MR, Drapeau MD, Yazdi PG, Shah KH, Moise DB, Thakar RR, Rauser CL, Mueller LD (2002) Evolution of late-life mortality in *Drosophila melanogaster*. Evolution 56:1982–1991

Rose MR, Rauser CL, Mueller LD (2005) Late life: a new frontier for physiology. Physiol Biochem Zool 78:869–878

Rose MR, Rauser CL, Mueller LD, Benford G (2006) A revolution for aging research. Biogerontology 7:269–277

Schoeni RF, Freedman VA, Martin LG (2008) Why is late-life disability declining? Milbank Q 86:47–89

Shahrestani P, Tran X, Mueller LD (2012a) Patterns of male fitness conform to predictions of evolutionary models of late-life. J Evol Biol 25:1060–1065

Shahrestani P, Tran X, Mueller LD (2012b) Physiological decline prior to death in *Drosophila melanogaster*. Biogerontology 13:537–545

Vaupel JW, Manton KG, Stallard E (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16:439–454

Vaupel JW, Carey JR, Christensen K, Johnson TE, Yashin AI, Holm NV, Iachine IA, Kannisto V, Khazaeli AA, Liedo P, Longo VD, Zeng Y, Manton KG, Curtsinger JW (1998) Biodemographic trajectories of longevity. Science 280:855–860

Verbrugge LM, Jette AM (1994) The disablement process. Soc Sci Med 38:1–14

Wax TM, Goodrick CL (1978) Nearness to death and wheel-running behavior in mice. Exp Geront 13:233–236

Yang J, Tower J (2009) Expression of hsp22 and hsp70 transgenes is partially predictive of *Drosophila* survival under normal and stress conditions. J Gerontol 8:828–838

Zhang WB, Pincus Z (2016) Predicting all-cause mortality from basic physiology in the Framingham heart study. Aging Cell 15:39–48