Published in final edited form as:

Exp Gerontol. 2012 October ; 47(10): 787–791. doi:10.1016/j.exger.2012.06.012.

Graphical and Demographic Synopsis of the Captive Cohort Method for Estimating Population Age Structure in the Wild

James R. Carey^{1,*}, Hans-Georg Müller², Jane-Ling Wang², Nikos Papadopoulos³, Alexis Diamantidis³, and Nikos Kouloussis⁴

¹Department of Entomology, University of California, Davis, California, 95616 U.S.A.

²Department of Statistics, University of California, Davis, California, 95616 U.S.A.

³Department of Agriculture Crop Production and Rural Environment, University of Thessaly, Phytokoy St. 38446 N. Ionia (Volos), Magnisias, GREECE

⁴Laboratory of Applied Zoology and Parasitology, School of Agriculture, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

Abstract

The purpose of this paper is to complement the literature concerned with the captive cohort method for estimating age structure including (1) graphic techniques to visualize and thus better understand the underlying life table identity in which the age structure of a stationary population equals the time-to-death distribution of the individuals within it; (2) re-derive the basic model for estimating age structure in non-stationary population in demographic rather than statistical notation; and (3) describe a simplified method for estimating changes in the mean age of a wild population.

Keywords

life table; stationary population; age class; insect demography; death distribution

INTRODUCTION

First introduced to the literature by Müller and his colleagues (Müller et al., 2004), the captive cohort method uses data gathered on the remaining lifespans of captured individuals to estimate the age structure of the wild population from which they are sampled. The underlying concept is that a population's age structure and its death distribution are inextricably and uniquely interconnected. Therefore with certain simplifying assumptions the former can be estimated from the latter using demographic models and reference life

© 2012 Elsevier Inc. All rights reserved.

*Correspondence to: Department of Entomology One Shields Ave. University of California Davis, CA 95616 USA 530-752-6217 jrcarey@ucdavis.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

tables. Publications subsequent to the Müller et al paper (Müller et al., 2004) extended the theoretical foundations for the captive cohort method (Müller et al., 2007; Vaupel, 2009), estimated the age structure in a wild medfly population (Carey, 2011; Carey et al., 2008), explored the implications of post-capture patterns of reproduction (Kouloussis et al., 2011) and age bias in sampling (Kouloussis et al., 2009), and modified the concept to estimate the shifts in mean age (Carey et al., 2012).

We believe that the basic analytical foundations for the capture cohort method described in these articles (recent summary in Carey, 2011) can be enhanced in two respects, both of which serve as the main focus of the current paper. First, by incorporating selected graphics designed to aid readers in visualizing and thus better understanding the concepts. Second, by converting the modelling notation from statistical to demographic in order to better situate it in the mainstream demographic literature and thus make it more accessible to a wider readership.

VISUALIZING THE LIFE TABLE IDENTIFY

Connecting Population Age-structure and its Death Distribution

A simple table (see Table 1 in Müller et al., 2004) illustrates the relationship between age distribution and remaining lifetime distribution for a stationary population (i.e. the identity). Table 1 shows the life table identity in the two bolded columns. The source of this identity is a simple arithmetic relation—the number of deaths in each time-to-death class equals the number of living in each population age class.

The graphs presented in Fig. 1 help visualize how these are connected. The upper panel in this figure depicts the percentage of the hypothetical population in each of four age classes and corresponds to the “proportion living” column in Table 1. The lower panel shows the time-to-death distribution in this hypothetical population and corresponds to the “proportion of deaths” column in this table. The bar segments labelled a, b, c and d in the top panel represent the respective percentages of each age class that die in time intervals 0-to-1, 1-to-2, 2-to-3, and beyond 3. All segments (a, b, c and d) are represented in the age class 0 bar (top panel) because this bar depicts members of the youngest population age class. Therefore these members will be subject to the risk of dying in *every* post-capture time interval. In contrast, all of the oldest members of the population in the top panel (age class 3) are contained in a single segment labelled “a”. This is because all of these members will die in the very first post-capture time interval.

All members in each of the age classes will die according to the mortality rates associated with each of their ages. The total deaths in the first time step corresponds to the sum of the segments labelled “a”. This result illustrates the general concept that applies to all post-capture time steps—the cumulative percentage of deaths in the first post-capture time interval (40%) equals the percentage of the original population in the first population age class (40%). Similarly the cumulative percentage of deaths from each of the remaining age classes shown in the bottom panel (i.e. 30, 25, and 5%, respectively) that occur over the subsequent time steps equals the percentage of the original population in the corresponding age classes (i.e. 30, 25, and 5%, respectively) shown in the top panel.

Although hypothetical, these examples illustrate the relationship of population age structure and the death distribution of the individuals from which it is constituted. In its basic and simplest version, the captive cohort concept directly utilizes the above mentioned life table identity. Assuming (A) that the captive cohort is randomly sampled from the wild cohort, so that each age group is sampled with probability corresponding to the frequency of this age group in the wild, and (B) that the survival schedule of the captive cohort remains the same as the survival schedule of the wild cohort, then (C) aging and survival in the wild can be inferred using the life table identity.

The identity behind the captive cohort method described in the earlier papers by Müller, Carey and their colleagues is the classical formula for residual renewal time in a stationary renewal process described in the early work on stationary population theory by both Lotka (1907, 1928) and Feller (1950). More recent work involving stationary population theory relevant to this method includes the classic papers by Ryder (1973, 1975) on replacement populations, the article by Kim and Aron (1989) showing the equivalency of the average age and average expectation of remaining life in a stationary population, the book section in Keyfitz (1985; p74) containing a general formula for the average expectation of life in a stationary population, the demography text Preston and co-authors (2001; pp53-58) outlining the basic properties of a stationary population, and the paper by Goldstein (2009) showing the equivalency of life lived and life left in stationary populations. Vaupel (2009) re-derived the basic model contained in Müller et al. (2004). He also shows that the captive cohort concept is a general model in which the proportion of individuals younger than a equals the proportion whose remaining lifespan is less than a . Or conversely, the proportion of individuals a or older is equal to the proportion of individuals who will still be alive in a years.

Pre- and Post-capture Life Segments: Medfly Example

The graphical perspective presented in this section (Fig. 2a,b) pertains to a hypothetical life table population of the medfly consisting of over 100 age classes that is used to illustrate two additional demographic concepts related to the captive cohort technique; i.e. the pre- and post-capture histories of individuals. Note from Fig. 2a that, whereas the majority of early post capture deaths (right-panel) were long-lived (older) individuals captured late in their lives, most of the last post-capture deaths were long-lived individuals captured early in their lives. These patterns underlie a general principle that is useful in interpreting capture data (Carey, 2011; Carey et al., 2008)—the relative presence of short- versus long-lived individuals post-capture reveals the relative presence in the population of older versus younger individuals.

A second important principle can be visualized, first by detaching the left-hand (pre-capture) segments of individual flies from their right-hand counterparts, and then by ordering these segments top-to-bottom from shortest to longest (Fig. 2b). The result of this process reveals the equivalency of pre- and post-capture life-days lived in a stationary population. This figure illustrates one of the central conclusions reached by James Vaupel in his paper related to the captive cohort method—the number of days a randomly-selected individual in a

stationary population has lived equals, on average, the number of days that individual has remaining (Vaupel, 2009).

DEMOGRAPHIC DERIVATION

The available formulae for the captive cohort method have been formulated in continuous time. Here we provide discrete time versions of these relationships that will facilitate a better demographic understanding of the connections between wild, captive and reference survival that underpin the concept of residual demography. We also facilitate use of these relationships when one has observations in the form of life tables as is commonly the case. By applying discrete time relationships one may avoid the step to convert such data into continuous hazard rates, which requires careful implementation (Müller et al., 1997; Wang et al., 1998).

The continuous time case of residual demography has been investigated by Müller et al (2004) for a simplified case with only a captive cohort and in Müller et al (2007) for a more general situation, where one has data for both a captive and a reference cohort and assumption (B) above, requiring the equality of the survival schedules in wild and captive cohorts is not needed. We focus here on the model that was proposed in Müller et al (2007) and was applied to field experiments with medflies in Carey et al (2008).

In the following, indices C, R, W refer to (population) quantities associated with the wild-caught “captive” cohorts (with individuals sampled from the wild population at various ages), reference cohorts that are age-homogeneous and are raised in the laboratory from birth, and wild cohorts, respectively. Here we use the demographic notation l_x to denote survival to age x and c_x to denote the fraction of the population at age x . Thus, $\sum_x c_x = 1$. Strictly speaking, the equations are approximations that hold for the population life tables. We do not distinguish between population and cohort (observed) life tables. Ages x and times x^* in the discrete formulas refer to the corresponding life table intervals; ideally they would correspond to the midpoints of these intervals. For applications to medflies the life table intervals are all of length one day.

Specifically, x^* denotes the time elapsed since capture of the captive cohort; since the ages of individuals in this cohort are unknown, x^* does not correspond to any individual's actual age, but is known, while individuals ages at capture and thus the ages of individuals in the captive cohort are unknown. Ages x stand for actual ages of individuals in the wild, the population distribution of which we aim to infer. With these conventions, eq. (3) in Müller et al (2007), which provides a basic relationship between the survival function of the captive population, the survival function of the reference cohort and the density of the age distribution of the wild population, is found to have the following discrete counterpart in demographic notation:

$$l_S(x^*) = \sum_x l_R(x+x^*) \frac{c_W(x)}{l_R(x)} \quad (1)$$

This is obtained from the continuous version by replacing the survival functions of the reference and captive cohorts by their discrete counterparts l_x , l_R and replacing the (continuous) age density in the wild' by c_W .

We note here that eq. (3) in Müller et al (2007), from which eq. (2) is derived, itself is a consequence of a formula for the residual renewal time in a stationary renewal process (Feller, 1950). A reviewer kindly pointed out the close connections with the renewal theory presented in Feller (1966; pp 353-356; 371), with some related precursor work described by Goldstein (2009).

Taking differences in eq. (1) simultaneously on both sides of the equation (corresponding to differentiation of both sides of eq. (3) in Müller et al (2007) one obtains the relation

$$c_S(x^*) = \sum_x c_R(x+x^*) \frac{c_W(x)}{l_R(x)} \quad (2)$$

Since the quantities c_S and c_R can be determined from the observed survival in captive and reference cohorts, eq. (2) may be used to determine the unknown target quantity c_W by numerical deconvolution. This deconvolution is not entirely straightforward, as the quantity of interest c_W is nested in the sum on the right hand side.

To ensure that the convolution equation can be solved for c_W requires certain identifiability assumptions such as decreasing wild age frequency distribution. Implementation requires a somewhat complex numerical regularization procedure. The latter can be implemented in various forms. As mentioned in Müller et al (2007), assumptions for the validity of the relationships (1), (2) are that (A') no age bias occurs in wild captures, when assembling the captive cohort, which is the same as Assumption (A) above, and (B') the force of mortality for any subject depends only on current but not past external conditions, i.e., the force of mortality is memoryless. We note that (B') is much weaker than the previous assumption (B). It is this weaker condition that requires additional measurements on a reference cohort, which is reared under laboratory conditions throughout.

An additional assumption for all captive cohort methods for inferring the survival schedule of a wild population is also that the wild population is stable and stationary. If this assumption holds, knowledge of c_W will lead to the wild survival schedule, as then

$$l_W(x) = \frac{c_W(x)}{c_W(0)} \quad (3)$$

holds. Violations of this assumption can be addressed with more complex modelling approaches, which require that one quantifies population birth rates or immigration/emigration rates (which will often be unrealistic).

A special case of the above considerations occurs if it may be assumed that

$$l_W = l_R \quad (3)$$

This is a relatively strong assumption that corresponds to (B) above and that was made by Müller et al (2004). This assumption has been used occasionally in previous applications of renewal models in industrial reliability (see references in Müller et al., 2004; Müller et al., 2007). This assumption was removed in the more general approach Müller et al (2007), to alleviate the strong restriction it imposes by replacing assumption (B) with the much weaker assumption (B').

If one is willing to make assumption (4), the relationships become considerably simpler and lead to

$$l_w(x) = \frac{c_c(x)}{c_c(0)} \quad (5)$$

so that one has an explicit solution that depends only on the observed age distribution in the captive cohort, thus eliminating the need for a reference cohort.

MEAN POPULATION AGE

Many biologists who study aging in wild insect populations do not have access to the sophisticated computer algorithms required to estimate the entire age distribution of wild populations derived in the previous section. A simplified approach uses regression that focuses on changes in mean population age rather than on the entire age distribution (Carey et al., 2012). The latter approaches use two demographic principles useful for interpreting post-capture longevity data. The first is the *reciprocity principle*: changes in the post-capture time-to-death are inversely related to the changes in the mean age of the population—e.g. an increase in mean post-capture age of death implies a decrease in the mean age of the wild population. The second the *proportionality principle*: decreases in the average post-capture times to death are associated with an opposite and approximately equal (i.e. $0.9 \approx 1.0$) increase in population age. The dashed lines in Fig. 4 illustrate these relationships. For example, the 30-day reductions in the post-capture longevity from 110 to 80 and again from 80 to 50 days imply roughly equivalent increases in the mean age of the population—from 5 to 38 day means (i.e. 33 day increase) and from 38 to 70 day means (i.e. 32 day increase).

Although not depicted here, Carey and his co-workers (Carey et al., 2012) noted that the post-capture extremes of shortest- and longest-lived individuals also provided important complementary information on age structure in the wild. They noted (A) that the presence of long-lived individuals in the post capture cohort must reveal the presence of recently-emerged individuals in the population at the time of sampling and, therefore, can be used as an indicator of dates of first and last emergence, and (B) that a downward trend in the post-capture longevity of the shortest lived individuals in population samples must reveal either increasing levels of population frailty or increasing fractions of the frail in the population.

DISCUSSION

The outcome of the processes involved in conceptualization, development, refinement, and extension of the captive cohort method is an example of how new concepts and approaches can emerge when researchers seek new approaches to practical problems, in this case

estimating population age structure in wild insect populations. The challenge in the current case was to find a method for estimating age structure in insect populations consisting of individuals of unknown age as a less costly and more efficient and accurate alternative to high tech approaches (Cook et al., 2006; Cook et al., 2008; Gerade et al., 2004; Krafur et al., 1995).

The captive cohort method has contributed to mainstream demography as well as to the biodemography of aging in the wild. (A) *Stationary population theory* (e.g. Goldstein, 2009; Preston et al., 2001; Ryder, 1975). The model on the capture cohort that was first introduced to the gerontology literature (Müller et al., 2004) but has a precedent in earlier statistical (Feller, 1950, 1966) and demographic literature (Keyfitz, 1985) revealed three equivalencies: (1) age structure and the distribution of remaining deaths; (2) mean age of those living and the mean age at which they all die; and (3) the probability of an individual who has lived x days and the probability this individual is one who has exactly x days remaining; (B) *Ecology of longevity*. One of the most important findings reported by Carey and his colleagues (Carey et al., 2008) in the study of medfly aging in the wild using captive cohorts was the extraordinary longevity potential in wild-caught flies relative to 'never-wild' flies of the same strain. These researchers noted that demographic selection would not easily account for the observed differences in mean remaining lifetimes between wild-caught and reference cohorts. This observation thus generated several new hypotheses for why flies that experience wild conditions have the potential for extended longevity in a protected environment including reproductive cost reductions, acquisition of health-promoting microorganisms, and a climatic tempering effect that increases hardiness.

CONCLUSIONS

Studies on aging, longevity and adaptation in the wild of model organisms have the potential to validate laboratory findings and generate new sets of questions on aging in a domain of aging research that complements the bench science of biogerontology (Carey, 2011). By complementing the earlier literature on the captive cohort method through the use of graphics that clarify the concept, notation that is consistent with that used in mainstream demography, and introducing a technique that focuses on mean population age and thus reduces the analytical requirement, we hope to encourage more research on organisms that are studied in evolutionarily-relevant (field) environments.

Acknowledgments

Research supported by NIA/NIH grants P01 AG022500-01 and P01 AG08761-10.

REFERENCES CITED

- Carey JR. Biodemography of the medfly: Aging, longevity, and adaptation in the wild. *Experimental Gerontology*. 2011; 46:404–411. [PubMed: 20933076]
- Carey JR, Papadopoulos N, Müller H-G, Katsoyannos B, Kouloussis N, Wang J-L, Wachter K, Yu W, Liedo P. Age structure changes and extraordinary life span in wild medfly populations. *Aging Cell*. 2008; 7:426–437. [PubMed: 18363903]

- Carey JR, Papadopoulos N, Papanastasiou S, Diamanditis A, Nakas CT. Inferring changes in population age using the death distributions derived from the daily capture of live medflies. *Ecological Entomology*. 2012 in press.
- Cook PE, Hugo LE, Iturbe-Ormaetxe I, Williams CR, Chenoweth SF, Ritchie SA, Ryan PA, Kay BH, Blows MW, O'Neil SL. The use of transcriptional profiles to predict adult mosquito age under field conditions. *Proceedings of the National Academy of Science*. 2006; 103:18060–18065.
- Cook PE, McMeniman CJ, O'Neil SL. Modifying insect population age structure to control vector-borne disease. *Advances in Experimental Medicine and Biology*. 2008; 627:126–140. [PubMed: 18510020]
- Feller, W. An introduction to probability theory and its applications. Wiley; New York: 1950.
- Feller, W. An introduction to probability theory and its applications. Vol. Vol. 2. Wiley; New York: 1966.
- Gerade BB, Lee SH, Scott TW, Edman JD, Harrington LC, Kitthawee S, Jones JW, Clark JM. Field validation of *Aedes aegypti* (Diptera: Culicidae) age estimation by analysis of cuticular hydrocarbons. *Journal of Medical Entomology*. 2004; 41:231–238. [PubMed: 15061283]
- Goldstein JR. Life lived equals life left in stationary populations. *Demographic Research*. 2009; 20:3–6.
- Keyfitz, N. Applied Mathematical Demography. 2nd ed. Springer-Verlag; New York: 1985.
- Kim YJ, Aron JL. On the equality of average age and average expectation of remaining life in a stationary population. *SIAM review*. 1989; 31:110–113.
- Kouloussis N, Katsoyannos B, Müller H-G, Wang J-L, Su Y-R, Carey JR. Seasonal trends in *Ceratitidis capitata* reproductive potential derived from live-caught adult females in Greece. *Entomologia Experimentalis et Applicata*. 2011; 140:181–188. [PubMed: 22791908]
- Kouloussis NA, Papadopoulos NT, Müller H-G, Wang J-L, Mao M, Katsoyannos BI, Duyck P-F, Carey JR. Life table assay for assessing relative age bias in medfly capture methods. *Entomologia experimentalis et applicata*. 2009; 132:172–181. [PubMed: 22844133]
- Krafsur ES, Moon RD, Kim Y. Age structure and reproductive composition of summer *Musca autumnalis* (Diptera: Muscidae) populations estimated by pterin concentrations. *Journal of Medical Entomology*. 1995; 32:685–696. [PubMed: 7473624]
- Lotka AJ. The progeny of a population element. *Science*. 1907; 26:21–22. [PubMed: 17754777]
- Lotka AJ. The progeny of a population element. *American Journal of Hygiene*. 1928; 8:875–901.
- Müller HG, Wang J-L, Carey JR, Caswell-Chen EP, Chen C, Papadopoulos N, Yao F. Demographic window to aging in the wild: Constructing life tables and estimating survival functions from marked individuals of unknown age. *Aging Cell*. 2004; 3:125–131. [PubMed: 15153180]
- Müller HG, Wang J-L, Yu W, Delaigle A, Carey JR. Survival in the wild via residual demography. *Theoretical Population Biology*. 2007; 72:513–522. [PubMed: 17727909]
- Müller HG, Wang JL, Capra WB. From lifetables to hazard rates: The transformation approach. *Biometrika*. 1997; 84:881–892.
- Preston, SH.; Heuveline, P.; Guillot, M. Demography: Measuring and Modeling Population Processes. Blackwell Publishers; Malden: 2001.
- Ryder NB. Two cheers for ZPG. *Daedalus*. 1973; 102:45–62.
- Ryder NB. Notes on stationary populations. *Population Index*. 1975; 41:3–28.
- Vaupel JW. Life lived and left: Carey's equality. *Demographic Research*. 2009; 20:7–10.
- Wang J-L, Müller H-G, Capra WB. Analysis of oldest-old mortality: Lifetables revisited. *Annals of Statistics*. 1998; 26:126–163.

Highlights

- New graphs help readers visualize the life table identity
- Statistical notation in the main model is replaced with demographic notation
- A new simplified model is described that estimates mean population age

Fig. 1.

Graphic showing the equality of the age distribution in a stationary population (upper) and the distribution of times-to-death of its members (lower). The implied age-specific mortality rates (q_x) are $1/4$, $1/6$, $4/5$ and 1.00 for age classes 0, 1, 2, and 3, respectively. Each segment labeled a, b, c or d in the top bar graphs corresponds to the fraction of the age class of the population that dies in the intervals, 0, 1, 2, and 3 respectively (bottom bar graphs). For example, the percentage of the standing population that die in each segment labeled 'a' (top graph) sum to 40% in the time-to-death class 0 (bottom graphic). Or alternatively, as the population ages the number of individuals in segments a, b, c, and d in age class 0 (top graph) die off at times 0, 1, 2 and 3, respectively as shown in the time-to-death distribution (bottom graph).

Figs. 2a,b.

Pre- and post-capture patterns depicted for the horizontal life-lines of 100 hypothetical medflies. In the inset in Fig. 2a each tic in the horizontal life lines depict the ages in which the individual medflies are captured, the left and right portions representing the pre- and post-capture segments, respectively. The pre-capture life days of individuals are connected to the post-capture life-days in Fig. 2a. These segments are decoupled from the post-capture life days and rank ordered top-to-bottom from shortest to longest to show their equivalency in Fig. 2b.

Fig. 3.

Schematic depicting hypothetical age distributions of younger (Population A) and older (Population B) populations and their corresponding post-capture death distributions, labelled A' and B', respectively. Note that the average age of death of members in Population A is greater than the average age of death for the members in Population B because the members are more youthful on average in the former population (redrawn from Carey et al., 2012).

Fig. 4.
Mean post-capture longevity regressed on mean population age in the Mediterranean fruit fly (redrawn from Fig. 2 in Carey et al., 2012).

Table 1

Relationship of the age distribution (c_x) of a four age class stationary population and the time-to-death distribution of its members (d_x^*) (see Müller et al., 2004 for details).

Age or Time	Population		Captures	
Age (x) or Time-to Death (x*)	Survival	Proportion Living	Survival	Proportion of Deaths
x or x*	l_x	c_x	l_x^*	d_x^*
0	1.000	0.40	1.00	0.40
1	0.750	0.30	0.60	0.30
2	0.625	0.25	0.30	0.25
3	0.125	0.05	0.05	0.05
4	0.000	0.00	0.00	0.00