$g(t|\alpha) = 1/(a+1-\alpha)$, $0 \leqslant t \leqslant 1$, $a \leqslant \alpha \leqslant a+1$. Hence, the *average risk time in $A_{a,p}$* is given by

$$\int_a^{a+1} E(T_{A_{a,p}}|\alpha)\,\mathrm{d}\alpha = \int_a^{a+1} \left(\int_0^{a+1-\alpha} t \cdot g(t|\alpha)\,\mathrm{d}t \right) \mathrm{d}\alpha = \int_a^{a+1} \left(\int_0^{a+1-\alpha} t \cdot \frac{1}{(a+1-\alpha)}\,\mathrm{d}t \right) \mathrm{d}\alpha$$

$$= \int_0^1 \left(\int_0^{1-\alpha} t \cdot \frac{1}{(1-\alpha)}\,\mathrm{d}t \right) \mathrm{d}\alpha = \int_0^1 \frac{1-\alpha}{2}\,\mathrm{d}\alpha = \frac{1}{4} \text{ person-year}$$

Deceased in $B_{a+1,p}$. A person who entered $A_{a,p}$ at age α, $a \leqslant \alpha \leqslant a+1$, and died in $B_{a+1,p}$ has been under risk throughout $A_{a,p}$ and contributes a risk time of $(a+1-\alpha)$ person-year in $A_{a,p}$. Thus, the *average risk time in $A_{a,p}$* is given by

$$\int_a^{a+1} E(T_{A_{a,p}}|\alpha)\,\mathrm{d}\alpha = \int_a^{a+1} (a+1-\alpha)\,\mathrm{d}\alpha = \int_0^1 (1-\alpha)\,\mathrm{d}\alpha = \frac{1}{2} \text{ person-year}$$

The risk time of such a person in $B_{a+1,p}$ varies on $[0, \alpha - a]$ and is assumed to be equally distributed, giving the respective pdf as $g(t|\alpha) = 1/(\alpha - a)$, $0 \leqslant t \leqslant 1$, $a \leqslant \alpha \leqslant a+1$. Hence, the *average risk time in $B_{a+1,p}$* is estimated by

$$\int_a^{a+1} E(T_{A_{a,p}}|\alpha)\,\mathrm{d}\alpha = \int_a^{a+1} \left(\int_0^{\alpha-a} t \cdot g(t|\alpha)\,\mathrm{d}t \right) \mathrm{d}\alpha = \int_a^{a+1} \left(\int_0^{\alpha-a} t \cdot \frac{1}{\alpha - a}\,\mathrm{d}t \right) \mathrm{d}\alpha$$

$$= \int_0^1 \left(\int_0^{\alpha} t \cdot \frac{1}{\alpha}\,\mathrm{d}t \right) \mathrm{d}\alpha = \int_0^1 \frac{\alpha}{2}\,\mathrm{d}\alpha = \frac{1}{4} \text{ person-year}$$

REFERENCES

1. Carstensen B. Age–period–cohort models for the Lexis diagram. *Statistics in Medicine* 2007; **26**(15):3018–3045. DOI: 10.1002/sim.2764.
2. Osmond C, Gardner MJ. Age, period, cohort models. Non-overlapping cohorts don't resolve the identification problem. *American Journal of Epidemiology* 1989; **129**(1):31–35.
3. Sverdrup E. Statistiske metoder ved dødelikhetsundersøkelser (in Norwegian). *Statistical Memoirs*. Institute of Mathematics, University of Oslo, 1967.
4. Hoem JM. Fertility rates and reproduction rates in a probabilistic setting. *Biométrie-Praximétrie* 1969; **10**(1):38–66.
5. Hoem JM. Correction note. *Biométrie-Praximétrie* 1970; **11**(1):20.

Published online 11 September 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.3041

AUTHOR'S REPLY

Age–period–cohort models for the Lexis diagram, *Statistics in Medicine* 2007; **26**:3018–3045

I thank Joachim Rosenbauer and Klaus Strassburger (R&S) for their interest in the paper and the derivations of the population risk time [1].

1. ASSUMPTIONS ABOUT DISTRIBUTION OF DEATHS

The core of their argument is that computation of risk time in age $\times$ period $\times$ cohort subsets of a Lexis diagram should be done conditional to the age at entry into the upper triangle sets, termed A both in my paper and their letter.

R&S derive precisely what I in the paper only asserted by handwaving, namely that the contribution of risk time in the sets A and B from persons in age class a at calendar time p, who survive to time $p+1$ (and hence at that time in age class $a+1$), is on average $\frac{1}{2}$ year in each. Underlying the calculation is the perfectly reasonable assumption that the distribution of ages at time p among those surviving to time $p+1$ is uniform on the interval $[a, a+1]$—the integration is with respect to age at time p using the uniform measure on $[a, a+1]$. This assumption is only approximately the same as the assumption of a uniform age distribution for *all* persons alive at p ($L_{a,p}$) if mortality rates do not vary dramatically by age. The same arguments will hold in the case of computing the risk time for all those who die in the set $\mathsf{A} \cup \mathsf{B}$.

But R&S use exactly the same argument and assumptions when they compute the risk time separately in each of the sets A and B from persons who die in these sets, particularly the assumption of uniform age distribution at time p. However in the case of deaths this is highly untenable. Given that a person has died in the set A makes it much more likely that the person entered at an early age and had a long exposure.

Hence, the computation by R&S is correct but relies on different assumptions, which I consider counterintuitive.

2. A SMALL SIMULATION STUDY

To illustrate this I carried out a small simulation study as follows: R&S use the assumption that the age at time p for persons who die in either A or B is uniformly distributed over the interval $[a, a+1]$, and in order to complete the computations the additional assumption that the time of death given entry age α is uniformly distributed on $[p, p+1-\alpha]$. This is sufficient to simulate a number of deaths in each of A and B. Once this is done, the empirical distribution of ages at p and of the risk time in each of the sets can be computed. Similarly, the assumptions that I use in my paper, uniform distribution of deaths over $\mathsf{A} \cup \mathsf{B}$, are easily simulated and the same computations on the simulated sample carried out.

The difference between the two approaches is illustrated in Figure 1. The top part represents 800 deaths in each of A and B with ages at p uniformly distributed and deaths uniformly distributed within the possible follow-up (R&S assumptions). The lower half represents 1600 deaths uniformly distributed over A and B (my assumptions).

The assumptions that R&S make imply a very odd clumping of deaths in the corners of A and B, which is indeed difficult to find a justification for. The assumptions that I make in the paper induces a highly skewed distributions of age at p, given death in either A or B, which however is a perfectly sensible consequence.

The empirical means of the risk time for the deaths based on the simulated samples are shown too, demonstrating that the formulae derived by R&S and me are actually reproduced in the simulations.

R&S note that in my derivation persons who die in B contribute on average the same amount of risk time in A and B, which they find odd since anyone who dies in B has lived through A.

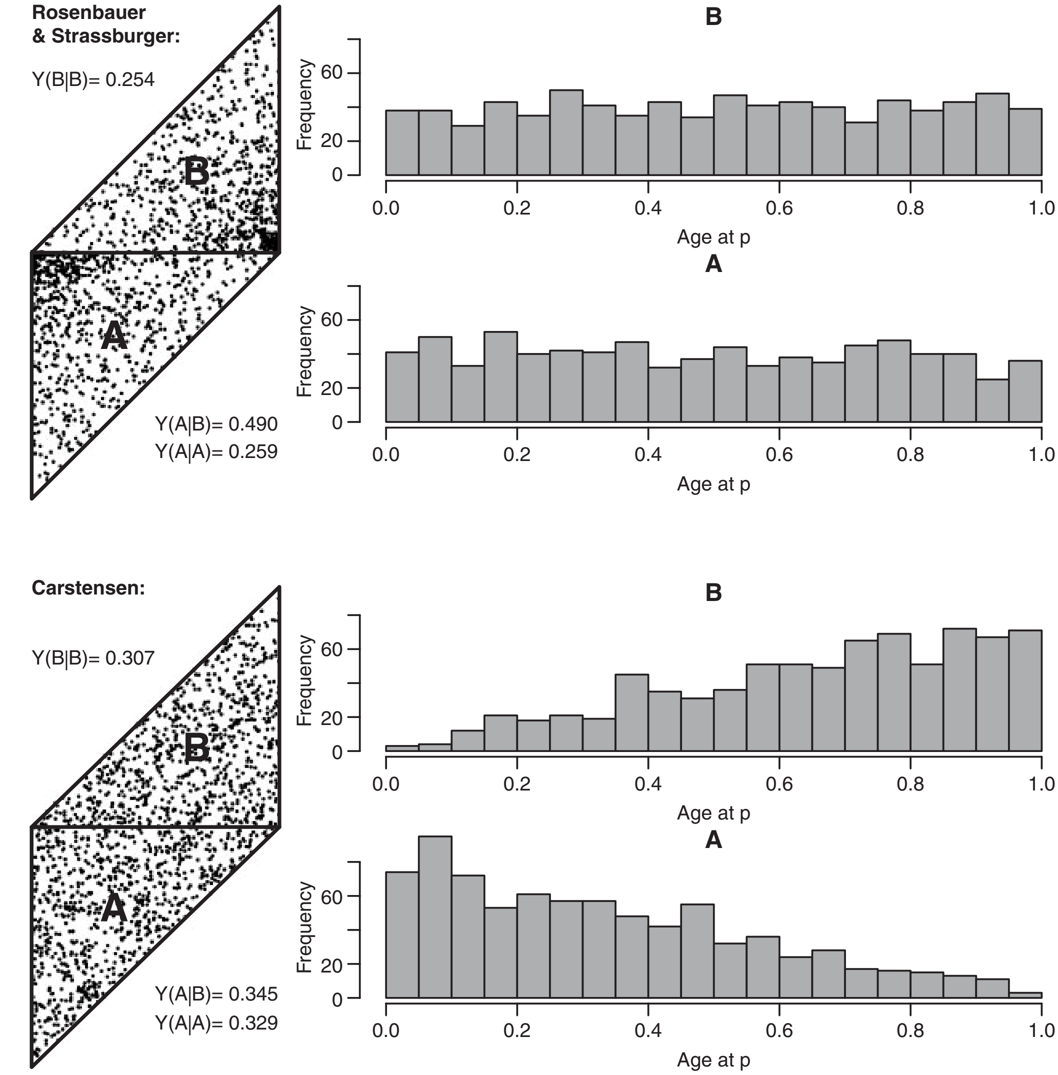

Figure 1. Results from simulation of deaths by two approaches. Top half: uniform distribution of age at start—a very strange assumption and not recommendable. Bottom half: uniform distribution of deaths over A and B—reasonable approximation in practise. The R program that does the simulation and the plot is available as http://staff.pubhealth.ku.dk/∼bxc/APC/R/Rosenbauer-Strassburger.R.

But not all persons dying in B have spent the same time in A; the shorter the time spent in A, the longer the time spent in B. That is the intuitive explanation—the mathematical one is in my paper.

3. CONCLUSION

My derivation in the paper [2] is correct and based on demographically sensible assumptions; the derivation by R&S is correct but based on assumptions that are highly unlikely to be relevant in any practical circumstances.

Therefore, the formulae given in my paper [2] are those that should be used in practice.

B. CARSTENSEN
Steno Diabetes Center, *Niels Steensens Vej 2*
DK.2820 Gentofte, *Denmark*
E-mail: bxc@steno.dk

REFERENCES

1. Rosenbauer J, Strassburger K. Letter: Re: Age–period–cohort models for the Lexis diagram. *Statistics in Medicine* 2007; **26**:3018–3045.
2. Carstensen B. Age–period–cohort models for the Lexis diagram. *Statistics in Medicine* 2007; **26**:3018–3045.

Published online 4 September 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.3058

