



Editor's choice

Sun exposure and longevity: a blunder involving immortal time

Jane E Ferrie* and Shah Ebrahim

*Corresponding author. E-mail: jane.ferrie@bristol.ac.uk

Unfortunately we have to start this Editor's Choice with an acknowledgment that we have fallen prey to a common, perennial problem; immortal time bias.

To illustrate the concept we borrow an example from William Farr, as used by James Hanley and Bethany Foster in a full and entertaining exposition of the problem in this issue of the journal.¹ Generals and bishops live longer than corporals and curates—but this is not necessarily because an elevated occupational status makes you live longer—it may simply be because you have to reach a certain age before it is possible to hold such positions. People become generals and bishops in middle age so their deaths arise after this point in time, whereas corporals and curates can die at any age above 20 or so.² This difference in time during which an event can occur to one group but not the other produces a bias favouring longer life expectancy—immortal time bias. In the figure on the next page, the problem is evident at a glance (Figure 1).³

In the October issue of the *International Journal of Epidemiology (IJE)* last year, we published a paper by Peter Brøndum-Jacobsen and colleagues in which they examined the effects of sunlight exposure on mortality among the whole population of Denmark aged above 40 years, using linked data from national registries.⁴ They used non-melanoma skin cancer as a proxy for sun exposure, which is a clever idea but it should have been obvious that the findings were 'too good to be true'—an apparent halving of all-cause mortality and reductions in myocardial infarction and hip fracture. The authors concluded: 'Causal conclusions cannot be made from our data. A beneficial effect of sun exposure per se needs to be examined in other studies'.

The Danish media picked up the story and it became front page news—'Sunbathers live longer'.⁵ Although the authors never made this claim in their published paper,

their interviews with the press did not appear to emphasize their non-causal conclusion. The Danish Cancer Association claims that this paper has undone all their good work in persuading Danes to keep out of the sun to avoid skin cancers.

Commentators on the story identified a likely problem of immortal time bias. People in the 'sun exposure' group had to live long enough to be diagnosed with skin cancer but the comparison group only had to be over 40 years old—the design of the study had built in a potential bias in favour of longevity among those presumed to be more highly exposed to sunlight. Theis Lange and Neils Keiding, in a letter commenting on the paper, pose questions about how such highly improbable findings got through the editorial process at *IJE*.⁶

In response to this criticism, Brøndum-Jacobsen and colleagues argue that their paper used both cohort and case-control analyses, and that the latter should be free from immortal time bias as cases and controls were matched on age.⁷ They acknowledge that the case-control analyses—which showed much smaller survival advantage [odds ratio (OR): 0.97, 95% confidence interval (CI) 0.96 to 0.99; vs hazard ratio (HR): 0.52, 95% CI 0.52 to 0.53]—should have been included in their abstract. In addition, they conducted a revised Cox proportional hazards analysis stratified by 10-year, 5-year and 2-year age strata in an attempt to control for immortal time bias, and interpret these findings as similar to those in their original paper. However, they fail to stress that the effect sizes become increasingly attenuated as the age matching becomes more exact, suggesting that the apparent effect of sun exposure may indeed be produced by immortal time bias.

Ironically, in parallel with the review and publication of this paper we had commissioned an 'Education Corner'

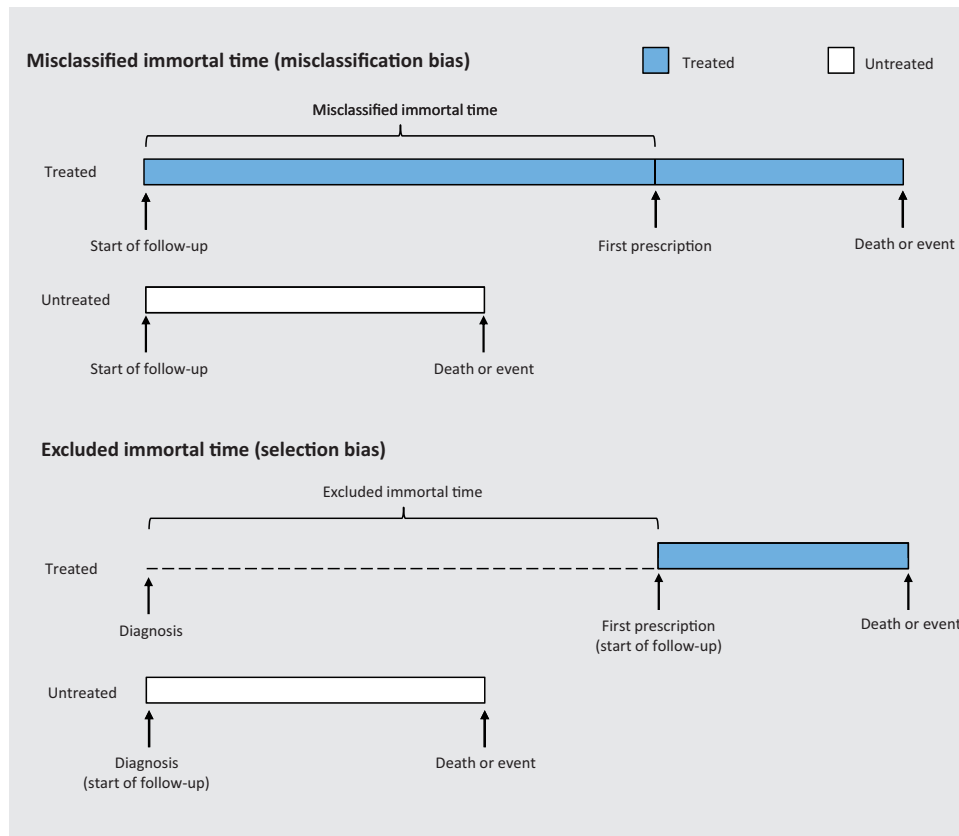


Figure 1. Immortal time bias is introduced in cohort studies when the period of immortal time is either incorrectly attributed to the treated group through a time fixed analysis (top) or excluded from the analysis because the start of follow-up for the treated group is defined by the start of treatment and is, by design, later than that for the untreated group (bottom). Reproduced with kind permission from the *British Medical Journal*.³

paper by Hanley and Foster on 'Avoiding blunders involving immortal time'. At the editors' request they added a postscript commenting on the Danish analyses.¹ Using a Danish population of over 4 million people drawn from the Human Mortality Database, they modelled the effect on all-cause mortality of an annual prize allocated at random. This mimics the incidence of non-melanoma skin cancer, but clearly the prize could have no biological effect on longevity. However, the analysis almost exactly mirrored both the original published findings and the revised age-strata analyses produced in response to Lange and Keiding's criticism. The effects reported by Brøndum-Jacobsen and colleagues could clearly be spurious.

Should the *IJE* have identified these flawed findings during the editorial process? The short answer is 'Yes' and our reviewers did indeed spot the problem: 'For the non-melanoma skin cancer group, you have to survive long enough to get non-melanoma skin cancer before you can die—i.e., you cannot die before the age of acquiring non-melanoma skin cancer'. In response to this and several other comments, the authors conducted a revised analysis excluding people under 40 years and applying different methods of analysis, and seemed to consider that by

truncating the age range they had dealt with the reviewer's comment above. Our reviewer considered the revised analysis to be an improvement and did not comment on the issue again. The paper was considered 'clever' and 'innovative' by our reviewers and was a large study apparently confirming earlier findings. The handling editor considered that the authors had done a sufficiently good job in dealing with the criticisms, and an editor-in-chief then accepted the paper for publication. The authors' matched case control analyses provided more plausible findings but we failed to ensure that these findings were given prominence or substituted for the misleading Kaplan-Meier and Cox model analyses.

Should this paper be retracted now? There are many examples of flawed analyses and inappropriate conclusions in the biomedical literature. Neither the authors nor the editors and reviewers who let such papers slip through the net are guilty of intentional mischief or fraud. We all learn from mistakes, and removing authorial and editorial mistakes from the public record is not a good solution. On the editorial side, like all who have fallen into this trap, we need to be more vigilant in the future. We have added a brief description of the problem and links to the material

in this issue, which in effect amount to ‘post-publication peer review’, to the online version of the paper by Brøndum-Jacobsen and colleagues.³ We believe that this editorial comment, the accompanying letters and Hanley and Foster’s excellent overview of immortal time bias provide a better understanding of the problem, how to detect it and how to deal with it properly.

Do Republican presidents kill babies?

Further methodological debate is provoked by a paper in this issue of the journal by Javier Rodriguez and colleagues.⁸ The provocative title—above—of a US blog on the paper⁹ reminded us that US presidents, Democrat and Republican, have frequently been indicted with killing babies and children in other countries: ‘Hey, hey LBJ. How many kids did you kill today?’ Rodriguez and colleagues look at the potential of presidents to contribute to infanticide at home.⁸ As the authors point out, infant mortality rates in one of the world’s wealthiest countries are shocking. Despite a dramatic downward trend between 1965 and 2010, for the period 2005–10 the rate exceeded that of most other developed countries as well as some less developed countries, like Cuba.¹⁰

Judging by its widespread acceptance, if not translation into effective policy, few quibble with the assertion by Sir Michael Marmot that the causes of the causes of inequalities in health reside in the social and economic arrangements of society.¹¹ Rodriguez and colleagues seem to have drawn particular flak for having located their study within the emerging sub-discipline of ‘political epidemiology’. In his many writings on inequalities, Marmot rarely strays into the overtly party-political. However, it is here that the causes of the causes can be at least partially addressed. Despite depressing similarities, in most countries there are real differences in health and social welfare policies between the main political parties. As Rodriguez and colleagues point out, it would be surprising were these not related to health outcomes, especially among vulnerable groups.¹²

To test this, Rodriguez and colleagues examined associations between the party of the last nine US presidents—four Democratic and five Republican—and infant mortality rates from 1965 to 2010. Their regression estimates show that, relative to trend, infant mortality rates during Republican administrations have been, on average, 3% higher than during Democratic administrations. These findings remained after adjustment for factors like unemployment, smoking, abortion rates, education and income.¹² The authors finish their paper on a cautionary note: ‘Further research is needed to determine whether the association we have uncovered is causal, and to identify

the mechanisms involved’. Coverage of the paper in the Washington Post is similarly cautious: ‘There is a correlation here that persists after accounting for some obvious alternative explanations. However, the mere existence of this correlation does not permit any strong conclusions’.¹³

In a commentary on the paper, Ralph Catalano takes Rodriguez and colleagues to task for providing so little in the way of explanatory mechanisms,¹⁴ although the authors do indicate possibilities, such as austerity vs increased social welfare in response to economic crisis.⁸ Danny Dorling in his commentary suggests psychosocial and behavioural as well as material mechanisms.¹⁵ However, Catalano’s main criticism is reserved for the methods. Using data from the Human Mortality Database he sets out to show that the findings of Rodriguez and colleagues are a product of their methods. In so doing he adds artefact to the potential explanations proffered and completes the quartet of potential explanations: material, psychosocial, behavioural and artefact, proposed by the Black Report on inequalities in health.¹⁶ Rodriguez and colleagues respond to Catalano with arguments about the relative merits of cubic spline and Box-Jenkins methods.¹² The editors of the *IJE* are of the opinion that there is no definitively ‘right’ answer for interpreting time trends, and so welcome this informative debate which will no doubt continue.

Methods of measurement and the mirror test

‘If you really want to know whether you are obese, just undress and look at yourself in the mirror’. This, according to Henry Blackburn and David Jacobs,¹⁷ was the advice given by Ancel Keys to participants visiting his laboratory who wanted to know if they were too fat. Undoubtedly the best possible indicator at the individual level, at the population level the mirror test is of more limited use. However, despite his reported fatphobe attitudes,¹⁷ it was Keys and colleagues who gave the ratio weight/height squared its now familiar name ‘body mass index’ and, in a comprehensive comparison of various measures of relative weight, endorsed it as the optimum obesity index.¹⁸

In 2010 a new *IJE* series was launched with an editorial by Debbie Lawlor and Nish Chaturvedi. The aim of the new series—‘Methods of measurement in epidemiology’—was to help ‘population health scientists to make informed decisions about the best measurement tools to use in different contexts and to understand the impact of using any one measurement tool’.¹⁹ Included in their editorial was an example of an area which the editors felt would benefit from inclusion in such a series—the measurement of body size and composition. Sadly, as yet, no one has risen to the challenge of addressing this issue. However, Keys’ original

paper and the accompanying three commentaries provide a comprehensive review of the merits and limitations of body mass index (BMI) as a measure of body fat and a risk factor for disease.^{17,20,21} Lawlor and Chaturvedi ponder the relative value of BMI compared with modern methods of measuring adiposity such as dual-energy X-ray absorptiometry scans, bioelectrical impedance and computed tomography or magnetic resonance imaging, particularly in children.¹⁸ The commentators give more prominence to alternative measures, cheaper and simpler to collect at population level, such as waist circumference and waist:hip ratio. However, time and again the discussion returns to the simplicity of BMI, its utility and its value as a marker of cardiometabolic risk.^{17–21}

Several of the empirical papers in this issue also focus on anthropometric measurements, mostly measures of adiposity, most often in relation to chronic conditions in later life. An elegant paper by Adam Hulman and colleagues presents the simultaneous effects of ageing and secular trends on the distribution of major cardiovascular risk factors in the UK over 25 years from 1985, using five phases of data from the Whitehall II study.²² In addition to blood pressure and lipids, the authors examined BMI and waist circumference in women and men aged 57–61. The bell curves for both measures flattened over successive phases. Smaller shifts to the left than to the right (higher values) indicate that most weight gain was seen among those already overweight and obese. However, although BMI changed little in the lean they still increased in girth. The authors suggest a simultaneous loss of muscle mass and accumulation of abdominal fat—a situation all too familiar to those of us in the requisite age group.

Emily Williams and colleagues used longitudinal data from a UK multi-ethnic population of older adults to examine associations between weight gain over two decades and disability.²³ They found both weight gain and moving up a BMI category to be associated with higher risks of three measures of later life disability: objectively measured locomotor dysfunction, self-reported functional limitation and problems with activities of daily living. Risks associated with weight gain were mitigated if accompanied by an increase in physical activity. However, highest levels of risk were observed among those who remained obese throughout.

While the potential for an association between weight gain and disability is immediately obvious, associations between BMI and autoimmune diseases are less so. In an 11-year follow-up of 75 000 women in the Danish National Birth Cohort, Maria Harpsøe and colleagues examined associations between pre-pregnancy BMI and 43 of the most common autoimmune diseases identified via national hospital in- and outpatient registers.²⁴ Risks of any autoimmune disease, dermatitis herpetiformis and type 1 diabetes increased with each unit increase in BMI, although the risks of celiac disease and Raynaud's phenomenon decreased. There were also higher risks of psoriasis, rheumatoid arthritis and sarcoidosis in obese compared with normal-weight women. The authors discuss potential explanations, including a common aetiology linking adiposity to autoimmunity, for example via changes in adipokine and cytokine levels, or shared risk factors, but also suggest that their novel findings need confirmation.

In addition to the potential risk of subsequent autoimmune disease, pre-pregnancy BMI has been associated with

Box 1. Daily intake of the overfed convicts³²

Breakfast	Dinner	Supper
12 oz. bread	16 oz. potatoes 16 oz. meat 6 oz. bread ½ oz. salt ¼ oz. pepper 1 pint soup	8 oz. bread
1 pint tea with ¾ oz. sugar & ⅙ oz. tea	1 oz. rice, barley or oatmeal for thickening	1 pint tea with ¾ oz. sugar & ⅙ oz. tea
Total weight	Total weight	Total weight
2 lb.	3 lb. 10 oz.	1 lb. 12 oz.

The total weight of food allowed per diem was thus seven pounds six ounces, including fifty-nine ounces of solid ailment; while to the blacksmiths and sawyers, an extra ration of 4 oz. of bread and 4 oz. meat was given, bringing their diet up to 7 lbs 14 oz., of which 67 oz. was solid ailment. (N.B. 67 oz. = 1.9 kilograms)

Box 2. Profile of a data resource 1856

A modern Magendie^a, wishing to investigate further into the subject of the food elements that build up, and the food elements that keep the animal fire alive, need sacrifice no more unhappy dogs at the shrine of Aesculapius^b. Let him but read in English blue books^c how, with an eye to the progress of physiological science, our authorities overfeed, underfeed, and don't feed at all, their convicts, their fighting men, and their paupers; and he will find sufficient facts ready made to afford him full occupation in arrangement and deduction for five years at least.³²

a. François Magendie (1783–1855), French physiologist

b. Greek god of medicine

c. The Blue Books — a series of British parliamentary and foreign policy documents published in blue cover since the seventeenth century

a number of adverse offspring outcomes including, more recently, general cognitive ability or intelligence.^{25–27} However, the problem with any study of associations between maternal BMI before pregnancy and offspring outcomes is the propensity for confounding by genetic factors and the postnatal environment. One approach to this problem is to compare associations between maternal BMI and offspring outcomes with those for paternal BMI.²⁸ A recent study which took this approach concluded that the similar association between maternal and paternal BMI and offspring intelligence suggests that it is not a specific pregnancy-related adiposity effect.²⁷ Another approach to minimizing the effects of familial factors is to use a sibling design.²⁹ This is one approach applied by Lisu Huang and colleagues in data from the Collaborative Perinatal Project.³⁰ The association they observed between maternal pre-pregnancy obesity and intelligence at age 7 years counter the findings of Mette Bliddell and colleagues²⁷ but confirm associations observed in earlier studies.^{25,26}

Confounding, criminality and overfed convicts

Intelligence is the main confounder of an observation by Amber Beckley and colleagues that short men are more likely to be convicted of violent crimes. The authors took advantage of Swedish register data for men who underwent military conscription tests between 1980 and 1992, to examine associations between height and first conviction for acts of violence such as homicide, assault and kidnapping. Over a mean time at risk of 27 years, just under 7% of the 713 877 conscripts were convicted.³¹ However, after adjustment for other anthropometric measures, socio-demographic factors and general cognitive ability, a weak but positive association between height and crime emerged. Muscle strength as well as height were measured

during the conscript examination and although, intuitively, we might expect stronger men to be more likely to engage in violence, the negative association between strength and conviction survived adjustment.

Nothing is mentioned of the crimes which brought 1554 convicts into the care of Dr Rennie, a medical officer in the penal colony at Fremantle, Western Australia, between 1 July and 31 December 1854.³² Almost all suffered from diseases of the skin, diseases of the digestive system or inflammatory eye disease: all, in Rennie's opinion, 'the results of overfeeding, assisted occasionally by a deficiency of vegetable matter'. Perusal of the daily diet of these men (Box 1) inclines to immediate endorsement of Rennie's view. Observing an absence of such disease in the general population, the efficacy of purgatives, and the cure effected by solitary confinement on a reduced diet of bread and water, Rennie suggests a reduction in the diet. Despite opposition from his superior, Rennie's suggestion was adopted and intake reduced from 59 to 46 ounces of solids per day, a course of action endorsed by reduced hospital admission rates 6 months later.

Throughout, the paper is littered with snide remarks and asides from the editor, e.g. 'The report of Dr Rennie met with some opposition from his superior, Dr Galbraith, who evidently has no leanings to commonsense deductions'. Later the editor asks the rhetorical question 'Why are convicts thus over replenished?' His answer is: 'The authorities find a body of men are more easily managed when well clad, well lodged, and supplied with more food than will satisfy their animal cravings'. David Cameron and other proponents of austerity worldwide should pause for thought although possibly, in their liaisons with Big Beverage, Big Food and Rupert Murdoch,³³ they feel they have their bread and circuses. Lastly, the editor of the *Journal of Public Health and Sanitary Review*, in an early forerunner of the *IJE's* Data Resource Profile series³⁴ describes a data resource for the observational epidemiologist of his time—Box 2.

References

- Hanley JA, Foster BJ. Avoiding blunders involving immortal time. *Int J Epidemiol* 2014;**43**:949–61.
- Farr W. *Vital Statistics: A Memorial Volume of Selections From the Reports and Writings of William Farr*. Humphreys NA (ed). London: The Sanitary Institute, 1885.
- Lévesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ* 2010;**340**:b5087.
- Brøndum-Jacobsen P, Nordestgaard BG, Nielsen SF, Benn M. Skin cancer as a marker of sun exposure associates with myocardial infarction, hip fracture and death from any cause. *Int J Epidemiol* 2013;**42**:1486–96.
- Politiken. [Sunbathers live longer.] Copenhagen: Politiken, 16 October 2013.
- Lange T, Keiding N. Skin cancer as a marker of sun exposure—a case of serious immortality bias. *Int J Epidemiol* 2014;**43**:971.
- Brøndum-Jacobsen P, Nordestgaard BG, Nielsen SF, Benn M. Authors' response to: Skin cancer as a marker of sun exposure—a case of serious immortality bias. *Int J Epidemiol* 2014;**43**:972–3.
- Rodriguez JM, Bound J, Geronimus AT. US infant mortality and the president's party. *Int J Epidemiol* 2014;**43**:818–26.
- Blattman C. Do Republican Presidents kill babies. <http://economistsview.typepad.com/economistsview/2013/11/infant-mortality-and-the-presidents-party.html> (17 April 2014, date last accessed).
- World Population Prospects: The 2012 Revision. United Nations, Department of Economic and Social Affairs, Population Division. 2013 <http://esa.un.org/unpd/wpp/Excel-Data/mortality.htm> (14 April 2014, date last accessed).
- British Medical Association. *Social Determinants of Health: What Doctors Can Do?* London: BMA, 2011.
- Rodriguez JM, Bound J, Geronimus AT. Rejoinder: Time series analysis and US infant mortality: de-trending the empirical from the polemical in political epidemiology. *Int J Epidemiol* 2014;**43**:831–34.
- Sides J. Do Republican presidents produce higher infant mortality rates? 2014. <http://www.washingtonpost.com/blogs/monkey-cage/wp/2014/02/14/do-republican-presidents-produce-higher-infant-mortality-rates/> (17 April 2014, date last accessed).
- Catalano R. Commentary: Political epidemiology, Republican presidents and dog food. *Int J Epidemiol* 2014;**43**:829–31.
- Dorling D. Commentary: All the presidents' children. *Int J Epidemiol* 2014;**43**:927–29.
- Blane D. An assessment of the Black Report's 'explanations of health inequalities'. *Sociol Health Illness* 1985;**7**:423–45.
- Blackburn H, Jacobs D Jr. Commentary: Origins and evolution of body mass index (BMI): continuing saga. *Int J Epidemiol* 2014;**43**:665–69.
- Kuys A, Fidanza F, Karvonen MJ, Kimura N, Taylor HL. Indices of relative weight and obesity. *Int J Epidemiol* 2014;**43**:655–69.
- Lawlor DA, Chaturvedi N. Methods of measurement in epidemiology—call for a new type of paper in the *IJE*. *Int J Epidemiol* 2010;**39**:1133–36.
- Wells JC. Commentary: The paradox of body mass index in obesity assessment: not a good index of adiposity, but not a bad index of cardio-metabolic risk. *Int J Epidemiol* 2014;**43**:672–74.
- Lukaski HC. Commentary: Body mass index persists as a sensible beginning to comprehensive risk assessment. *Int J Epidemiol* 2014;**43**:669–71.
- Hulmán A, Tabák AG, Nyári TA *et al*. Effect of secular trends on age-related trajectories of cardiovascular risk factors: the Whitehall II longitudinal study 1985–2009. *Int J Epidemiol* 2014;**43**:866–79.
- Williams ED, Eastwood SV, Tillin T, Hughes AD, Chaturvedi N. The effects of weight and physical activity change over 20 years on later-life objective and self-reported disability. *Int J Epidemiol* 2014;**43**:856–65.
- Harpsoe MC, Basit S, Andersson M *et al*. Body mass index and risk of autoimmune diseases: a study within the Danish National Birth Cohort. *Int J Epidemiol* 2014;**43**:843–55.
- Hinkle SN, Schieve LA, Stein AD, Swan DW, Ramakrishnan U, Sharma AJ. Associations between maternal prepregnancy body mass index and child neurodevelopment at 2 years of age. *Int J Obes (Lond)* 2012;**36**:1312–19.
- Tanda R, Salsberry PJ, Reagan PB, Fang MZ. The impact of prepregnancy obesity on children's cognitive test scores. *Matern Child Health J* 2013;**17**:222–29.
- Bliddal M, Olsen J, Størvring H *et al*. Maternal pre-pregnancy BMI and intelligence quotient (IQ) in 5-year-old children: a cohort based study. *PLoS One* 2014;**9**:e94498.
- Davey Smith G. Assessing intrauterine influences on offspring health outcomes: can epidemiological studies yield robust findings? *Basic Clin Pharmacol Toxicol* 2008;**102**:245–56.
- Donovan SJ, Susser E. Commentary: Advent of sibling designs. *Int J Epidemiol* 2011;**40**:345–49.
- Huang L, Yu X, Keim S, Li L, Zhang L, Zhang J. Maternal prepregnancy obesity and child neurodevelopment in the Collaborative Perinatal Project. *Int J Epidemiol* 2014;**43**:783–92.
- Beckley AL, Kuja-Halkola R, Lundholm L, Långström N, Frisell T. Association of height and violent criminality: results from a Swedish total population study. *Int J Epidemiol* 2014;**43**:835–42.
- Anon. Diseases of overfed convicts. *J Public Health and Sanitary Review* 1856;**2**:137–42.
- Ferrie JE. Progress, public health and vested interests. *Int J Epidemiol* 2013;**42**:1527–36.
- Lynch J, Stuckler D. 'In God we trust, all others (must) bring data'. *Int J Epidemiol* 2012;**41**:1503–06.



Letters to the Editor

Skin cancer as a marker of sun exposure: a case of serious immortality bias

From Theis Lange* and Niels Keiding

Department of Biostatistics, Institute of Public Health, University of Copenhagen, Denmark

*Corresponding author. Department of Biostatistics, Institute of Public Health, University of Copenhagen, Øster Farimagsgade 5, P.O.B. 2099, 1014 Copenhagen K, Denmark. E-mail: t.lange@biostat.ku.dk

Brøndum-Jacobsen *et al.* recently published in this journal¹ analyses of Danish register data concerning myocardial infarction, hip fracture and death from any cause, using incidence of skin cancer as indicator of high exposure to sunlight. The basic idea in the paper is that those who get a skin cancer diagnosis at any age are supposed to have been more exposed to the sun during their life than those who do not, and apparently the authors find it relevant to use ordinary prospective survival analysis to compare incidence of myocardial infarction, hip fracture and death from any cause between the two groups: those who (at some point) get a skin cancer diagnosis and those who do not.

Unfortunately, such an analysis is seriously flawed, because the definition of one of the two groups to be compared conditions on the future: in order to get a skin cancer diagnosis, and thus become a member of the skin cancer group, it is at least necessary to survive until age of diagnosis, but the authors' analysis does not take this conditioning into account. Put another way: for those in the skin cancer group it is impossible to die until the age of diagnosis of the cancer, the so-called immortal person-time.²

For ease of exposition we focus on the endpoint 'death from any cause'. It is seen in the lower left panel of Figure 2¹ that those who get non-melanoma skin cancer at some age have a hazard ratio of dying from any cause in the age interval 40–49 years of about 0.2 vs those who never get a non-melanoma skin cancer diagnosis. A main reason for

this is probably that very few of those with non-melanoma skin cancer are at all at risk for dying—most of the members of this group get their skin cancer diagnosis at ages >50 years and are therefore by design immortal in the age interval 40–49.

Methodology aside, we find it very surprising that neither the authors nor the editorial process have questioned the strange results at many places in the paper. For example: the upper right corner of Table 2¹ shows that persons who sooner or later get a diagnosis of malignant melanoma have a significantly reduced risk of dying from any cause: a hazard ratio of 0.89. Did no alarm bells sound? That the authors cautiously write 'causal conclusions cannot be made' in the abstract does not justify publishing a methodologically flawed analysis.

As a more comic point, we noted that *IJE* now quotes *P*-values with 308-digit precision—we hope that the chi-square approximation to the distribution of the log-rank statistic is justified!

References

1. Brøndum-Jacobsen P, Nordestgaard BG, Nielsen SF, Benn M. Skin cancer as a marker of sun exposure associates with myocardial infarction, hip fracture and death from any cause. *Int J Epidemiol* 2013;42:1486–96.
2. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd edn. Philadelphia, PA: Lippincott Williams & Wilkins, 2008.

Authors' Response to: Skin cancer as a marker of sun exposure—a case of serious immortality bias

International Journal of Epidemiology, 2014, 972–973

doi: 10.1093/ije/dyu102

Advance Access Publication Date:



From Peter Brøndum-Jacobsen,^{1,2} Børge G Nordestgaard,^{1,2} Sune F Nielsen¹ and Marianne Benn^{2,3*}

¹Department of Clinical Biochemistry, Herlev Hospital, Herlev, Denmark, ²Copenhagen University Hospital and Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark and ³Department of Clinical Biochemistry, Gentofte Hospital, Copenhagen, Denmark

*Corresponding author. Department of Clinical Biochemistry, Gentofte Hospital, Niels Andersenvej 65, DK-2900 Hellerup, Copenhagen, Denmark. E-mail: Marianne.Benn@regionh.dk

We thank Theis Lange and Niels Keiding for their interest in our report on the risk of skin cancer as a marker of sun exposure and risk of myocardial infarction, hip fracture and death from any cause.¹

Most studies are susceptible to certain biases, and inappropriate accounting of person-time in the design and analysis may in cohort studies introduce immortality bias. To address this and other potential biases, the data in our study were analysed both in a cohort design (prone to immortality bias) and in a case-control design, where each case was matched with five general population controls on the basis of age, birth year and gender (Tables 2 and 3 in the paper, respectively), and furthermore using both designs in age-strata of 10 years (Figures 2 and 3 in the paper, respectively).¹ In the matched case-control design, immortality bias is unlikely to be present, simply because both cases and controls had to be alive to the same age to be included for further follow-up. The directions of the risk estimates from the two different designs were similar, but effect sizes were attenuated in the matched case-control vs the cohort design, which is why we only concluded on the direction of risk estimates.

In Figure 1 below, we have now performed additional analyses in an attempt to exclude immortality bias using a modified approach. Within 10-year, 5-year and 2-year age-strata, we compared individuals diagnosed with non-melanoma skin cancer within a given age-stratum with those alive and without non-melanoma skin cancer in the same age-stratum. Importantly, those who develop non-melanoma skin cancer beyond the age-stratum enter into the analysis as not having non-melanoma skin cancer. We then followed these two groups for all-cause mortality within each of the age-strata shown in the figure. The results of the analyses are similar to those reported in the

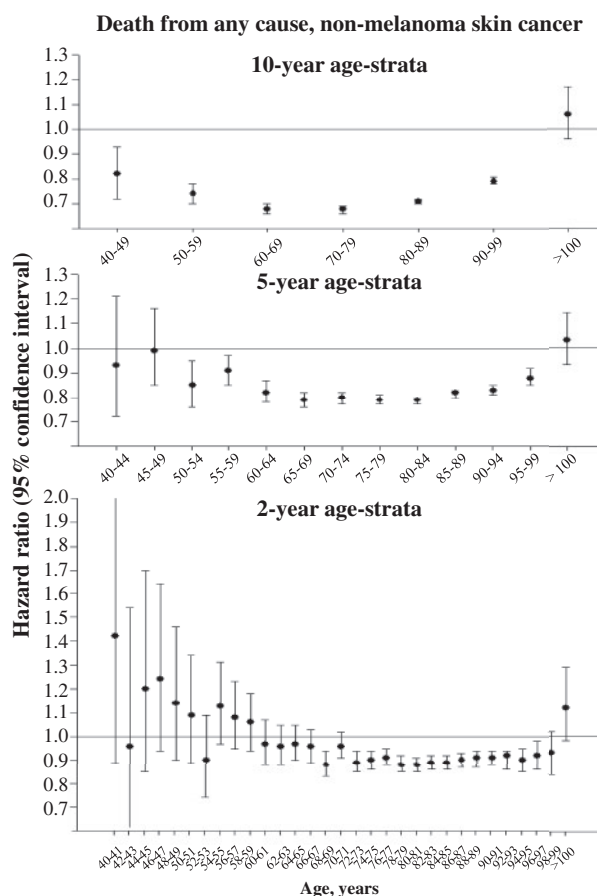


Figure 1. In the entire Danish population above 40 years of age, hazard ratios for death from any cause are shown within 10-year, 5-year and 2-year age-strata in individuals with vs without non-melanoma skin cancer. An individual with a non-melanoma skin cancer diagnosis occurring after the defined age-stratum was coded as an individual without a diagnosis of non-melanoma skin cancer.

paper, to us suggesting that non-melanoma skin cancer is associated with reduced death from any cause.

Interestingly, our results are in line with previous studies on non-melanoma skin cancer and all-cause mortality

in the Danish general population using similar databases: Jensen *et al.* have in two studies shown that individuals with basal cell carcinoma have reduced all-cause mortality.^{2,3}

We also thank Theis Lange for his contribution to the discussion of the interpretation of our results in the national Danish media. The national Danish newspaper *Politiken* reported our findings as the main front page story on 16 October 2013 under the headline 'Sunbathers live much longer'. The headline was decided exclusively by the newspaper, and we accepted the main text which in a simplified manner reported on our findings and even mentioned that no causal inference from sunbathing and skin cancer to myocardial infarction, hip fracture or death from any cause could be drawn from our study. An accompanying story focused on the Danish Cancer Society's yearlong advice to avoid being in the sun, due to the risk of skin cancer. A representative from the Society acknowledged that it is well known that those with skin cancer live longer. She also mentioned that the Society attributed longevity among those with skin cancer to more leisure time outdoor physical activity, rather than to positive effects from sunshine per se.

The Danish public, journalists at national TV and radio, and users of the internet etc. enjoyed the story, and soon almost everybody in Denmark knew that 'the more you are in the sun, the longer you live', a clear over-interpretation of our data. This is how stories sometimes develop in the media, beyond the scientist's control. Due to the high northern latitude of Denmark, Danes are deprived of sunshine for most of the year, and have for the past several years been told to stay away from it even when it is finally there. Therefore, many people in Denmark liked to be told that it was okay to be in the sun for a while, that is without the need to feel guilty.

As a consequence of this massive media attention, many prominent scientists in Denmark, including Theis Lange, read our paper and commented on its limitations (flaws, incorrect analyses etc.) in the media. In other words, our paper got a second round of revision after the one initially provided by the *International Journal of Epidemiology*. As science must often improve by peer review, we much appreciated this further review as well as the opportunity

now to respond to the letter by Theis Lange and Niels Keiding.

We are very cautious with respect to analyses and interpretation of national register data, and sincerely welcome advice on how to do this better in the future. Analyses will probably never be 'correct' and unequivocal. There are many possible pitfalls and potential biases, and careful thinking and many sensitivity analyses are often necessary when dealing with such data, as in present and previous studies.⁴

Rereading the paper, the results presented there in Table 3 and Figure 2, which are most likely unaffected by immortality bias, should have been presented in the abstract; however, we were restricted by a word count limit. Also, a discussion of immortality bias would have improved the paper and we are therefore happy to have this opportunity to address this. That said, we believe that the totality of data presented support the conclusion of our paper, which is that having a diagnosis of skin cancer is associated with less myocardial infarction, less hip fracture in those below age 90 years and less death from any cause, as the analyses not prone to immortality bias also support these conclusions.

Funding

This work was supported by the Danish Heart Foundation [10-01-R79-A2793-22574], the Faculty of Health Sciences, University of Copenhagen, and by Herlev Hospital, Copenhagen University Hospital.

References

1. Brondum-Jacobsen P, Nordestgaard BG, Nielsen SF, Benn M. Skin cancer as a marker of sun exposure associates with myocardial infarction, hip fracture and death from any cause. *Int J Epidemiol* 2013;**42**:1486–96.
2. Jensen AO, Bautz A, Olesen AB, Karagas MR, Sorensen HT, Friis S. Mortality in Danish patients with nonmelanoma skin cancer, 1978–2001. *Br J Dermatol* 2008;**159**:419–25.
3. Jensen AO, Lamberg AL, Jacobsen JB, Braae OA, Sorensen HT. Non-melanoma skin cancer and ten-year all-cause mortality: a population-based cohort study. *Acta Derm Venereol* 2010;**90**:362–67.
4. Nielsen SF, Nordestgaard BG, Bojesen SE. Statin use and reduced cancer-related mortality. *N Engl J Med* 2013;**368**:576–77.



Education corner

Avoiding blunders involving ‘immortal time’

James A Hanley¹ and Bethany J Foster^{1,2*}

¹Department of Epidemiology, Biostatistics, and Occupational Health and ²Department of Pediatrics, Montreal Children’s Hospital, Faculty of Medicine, McGill University, Montreal, QC, Canada

*Corresponding author. Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, 1020 Pine Avenue West, Montreal, Quebec, H3A 1A2, Canada. E-mail: james.hanley@mcgill.ca

Accepted 2 April 2014

As Groucho Marx once said ‘Getting older is no problem. You just have to live long enough’.
(Queen Elizabeth II, at her 80th birthday celebration in 2006)

This award proves one thing: that if you stay in the business long enough and if you can get to be old enough, you get to be new again.
(George Burns, on receiving an Oscar, at age 80, in 1996)
(Richard Burton died, a nominee 6 times, but sans Oscar, at 59. Burns lived to 100, so how much of the 41 years’ longevity difference should we credit to Burns’ winning the Oscar?)

Some time ago, while conducting research on U.S. presidents, I noticed that those who became president at earlier ages tended to die younger. This informal observation led me to scattered sources that provided occasional empirical parallels and some possibilities for the theoretical underpinning of what I have come to call the precocity-longevity hypothesis. Simply stated, the hypothesis is that those who reach career peaks earlier tend to have shorter lives.
(Stewart JH McCann. *Personality and Social Psychology Bulletin* 2001;27:1429–39)

Statin use in type 2 diabetes mellitus is associated with a delay in starting insulin.
(Yee *et al.* *Diabet Med* 2004;21:962–67)

Introduction

For almost two centuries, teachers have warned against errors involving what is now called ‘immortal time.’

Despite the warnings, and many examples of how to proceed correctly, this type of blunder continues to be made in a widening range of investigations. In some instances, the consequences of the error are less serious, but in others the false evidence has been used to support theories for social inequalities; to promote greater use of pharmaceuticals, medical procedures and medical practices; and to minimize occupational hazards.

We use a recent example to introduce this error. We then discuss: (i) other names for it, how old it is and who tried to warn against it; (ii) how to recognize it, and why it continues to trap researchers; and (iii) some statistical ways of dealing with denominators measured in units of time rather than in numbers of persons.

Example and commentary

Example

Patients whose kidney transplants (allografts) have failed must return to long-term dialysis. But should the failed allograft be removed or left in? To learn whether its removal ‘affects survival’, researchers¹ used the US Renal Data System to study ‘a large, representative cohort of [10 951] patients returning to dialysis after failed kidney transplant’. Some 1106, i.e. 32% of the 3451 in the allograft nephrectomy group, and 2679, i.e. 36% of the 7500 in the non-nephrectomy group, were identified as having died by the end of follow-up.

Patients in the two groups differed in many characteristics: to take into account a ‘possible treatment selection bias’, the authors constructed a propensity score for the

likelihood of receiving nephrectomy during the follow-up. They used this together with other potential confounders to perform ‘multivariable extended Cox regression’. The main finding of these analyses was that ‘receiving an allograft nephrectomy was associated with a 32% lower adjusted relative risk for all-cause death (adjusted hazard ratio 0.68; 95% confidence interval 0.63 to 0.74)’.

In their discussion, the researchers suggest that their findings of ‘improved survival’ after allograft nephrectomy ‘challenge the traditional practice of retaining renal allografts after transplant failure’. The title of the article (‘Transplant nephrectomy improves survival following a failed allograft’) suggested causality. They emphasized the large representative sample and the extensive and sophisticated multivariable analyses, but they did caution that ‘as an observational study of clinical practice, their analysis remains susceptible to the effects of residual confounding and treatment selection bias’ and that ‘their results should be viewed in light of these methodologic limitations inherent to registry studies’. They suggested that a randomized trial to evaluate the intervention in an unbiased way would be appropriate. Similar concerns about residual confounding and selection bias, and the need for caution, were expressed in the accompanying editorial reiterating the limitations of the ‘retrospective interrogation of a database’.

Commentary

‘Residual confounding’ may be a threat, but both authors and editorialists overlooked a key aspect of the analysis, one that substantially distorted the comparison. The overlooked information is to be found in the statements that:

3451 received nephrectomy of the transplanted kidney during follow-up; the median time between return to dialysis [the time zero in the Cox regression] and nephrectomy was 1.66 yr (interquartile range 0.73 to 3.02 yr).

(Paragraph 1 of Results section)

and that:

Overall, the mean follow-up was (only) 2.93 ± 2.26 yr.

(Paragraph 3 of Results section)

From these and other statements in the report it would appear that, in their analyses, follow-up of both ‘groups’ began at the time of return to dialysis. The use of this time-zero for the 3451 who had the failed allograft removed is not appropriate—or logical. These patients could not benefit from its removal until after it had been removed; but, as the median of 1.66 years indicates, a large portion of their ‘follow-up’ was spent in the initial ‘failed graft still in place’ state—along with those who never underwent nephrectomy of their failed allograft.

Since the 3451 patients who ultimately underwent a nephrectomy (the ‘nephrectomy group’) had to survive long enough to do so (collectively, approximately 6700 patient-years, based on the reported quartiles of 0.73, 1.66 and 3.02 years), there were, by definition, no deaths in these 6700 pre-nephrectomy patient-years. In modern parlance, these 6700 patient-years were ‘immortal’. There was no corresponding ‘immortality’ requirement for entry into the ‘non-nephrectomy group’. Indeed, all 10 951 patients returning to dialysis after failed kidney transplant began follow-up with their ‘failed graft in place’. Some 7500 of these remained in that initial state until their death (for some, death occurred quite soon, before removal could even be contemplated) or the end of follow-up, whereas the other 3451 spent some of their follow-up time in that initial state and then changed to the ‘failed graft no longer in place’, i.e. post-nephrectomy, state.

How big a distortion could the misallocation of these 6700 patient-years produce? The article does not have sufficient information to re-create the analyses exactly. [Figures 1 and 2](#) show a simpler hypothetical dataset which we constructed to match the reported summary statistics quite closely. It was created assuming no variation in mortality rates over years of follow-up or between those lived in the two states. The ‘virtual’ intervention was set up ‘retroactively’ and was limited to the dataset itself, rather than to real individuals, and so could not have affected (other than randomly) the mortality rates in the person-years lived in each state.

[Figure 2A](#) shows that even though the data were generated to produce the same mortality rate of 11.8 per 100 PY (person-years) in the person-years in the initial and post-‘intervention’ states, the inappropriate type of analysis used in the paper, applied to these hypothetical data, would have resulted in a much lower rate (6.4) in the ‘intervention group and a much higher one (17.1) in the ‘non-intervention’ group. The reason is that none of the 1031 deaths post-‘intervention’ could have occurred, and none of them did occur, in the 6732 (immortal) pre-‘intervention’ PY that are included in the denominator input to the rate of 6.4: logically, the 1031 post-‘intervention’ deaths only occurred in the post-‘intervention’ PY. And conversely, the 2759 deaths occurred not in 16 096 PY, but rather in the much larger denominator of $16\,096 + 6732 = 22\,828$ PY lived in the initial state. The omission of the 6732 PY from the denominator input led to the rate, higher than it should have been, of 17.1 deaths/100 PY. Indeed it was because of these (misplaced) immortal 6732 PY they had already survived that the 3451 patients got to have the ‘intervention’; in other words, it may not have been that they lived longer because they underwent the ‘intervention’, but rather that they underwent the ‘intervention’ because they survived

Hypothetical lifelines constructed so that the observed mortality rate is 3785 deaths in (10,951 x 2.93 = 32,086) patient years (PY), i.e., 11.8 per 100PY (as in the actual nephrectomy study), but with no variation in the rate over years of follow-up, and no difference (other than random) in rates in the experiences in the states of the organ of concern ('in place' or 'removed').

We constructed the lifelines by distributing the numbers of new cohort entries in a smooth decreasing pattern over the 11 calendar years, and applying the death rate of 11.8 per 100PY to the various resulting lengths of available follow-up, until the total number of deaths matched the reported 3785 and the number of PY of follow-up matched the reported 32,086. The 10,951 hypothetical lifelines (3785 completed, 7166 censored) were then ordered from shortest to longest.

Finally, starting from the day of return to dialysis and working forward, each follow-up day a number of persons were chosen randomly from among those had not already been selected, were still alive, and being followed that day. These persons were designated to undergo the 'intervention'. The timings of the interventions (3471 in all) were adjusted until the median and quartiles of the delay between return to dialysis and the intervention matched those in the article.

The selections from the generated names were performed blindly in 2012, in a retroactive lottery, applied in a forward direction, beginning in January 1994, to lifelines that had already run up to December 2004. Just as in Leibovici, and in Turnbull et al., the interventions were limited to the computer file, and so could not have affected the comparative mortality rates.

Selected lifelines, ordered from shortest to longest

Timelines that end in a straight edge were ended by death

Timelines that end in an angled edge were censored

A change from the 'initial to the 'post-intervention' status is denoted by an 'x'

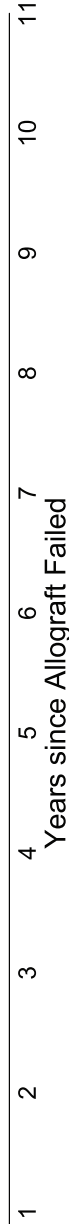


Figure 1. Excerpts from the simulated mortality experience in the contrasted ('organ intact' vs 'organ removed') states. Hypothetical lifelines were generated to have an average mortality rate of 3785 deaths in (10951 x 2.93 = 32 086) patient-years (PY), i.e., 11.8 per 100 PY (as in the actual nephrectomy study¹), but with no variation over years of follow-up, and no difference (other than random) between states ('name intact' or 'name removed'). We constructed the dataset by first generating names for 10 951 fictional persons, then distributing the numbers of new cohort entries in a smooth decreasing pattern over the 11 calendar years, and then applying the death rate of 11.8 per 100 PY to the various resulting lengths of available follow-up, until the total number of deaths matched the reported 3785 and the number of PY of follow-up matched the reported 32 086. The 10 951 hypothetical lifelines (3785 completed, 7166 censored) were then ordered from shortest to longest. Finally, starting from the day of return to dialysis and working forward, each follow-up day a number of persons were chosen randomly from among those who had not already been selected, were still alive and were being followed that day. These persons were designated to undergo an electronic 'removal' whereby, within the database, just their names (not their failed transplants) were (electronically rather than surgically) removed. The timings of these 'removals' (3471 in all) were set so that the median and quartiles of the delay between return to dialysis and becoming nameless matched the delays in the article. The selections, made by a random number generator in 2012, were made blindly, in a retroactive lottery, applied in a forward direction, beginning in January 1994, to lifelines that had already run up to December 2004. Just as in Leibovici¹⁶ and in Turnbull *et al.*,¹⁷ these interventions were limited to the 2012 computer file, and could not have affected the comparative mortality rates. Shown are 30 such lifelines selected systematically from these 10 951 ordered hypothetical ones, with a completed lifeline indicated by a single straight line, and a censored one by a pair of lines forming an arrowhead. The timing of the name removal is indicated by an x, and the post-intervention PY by red rather than grey boundary lines.

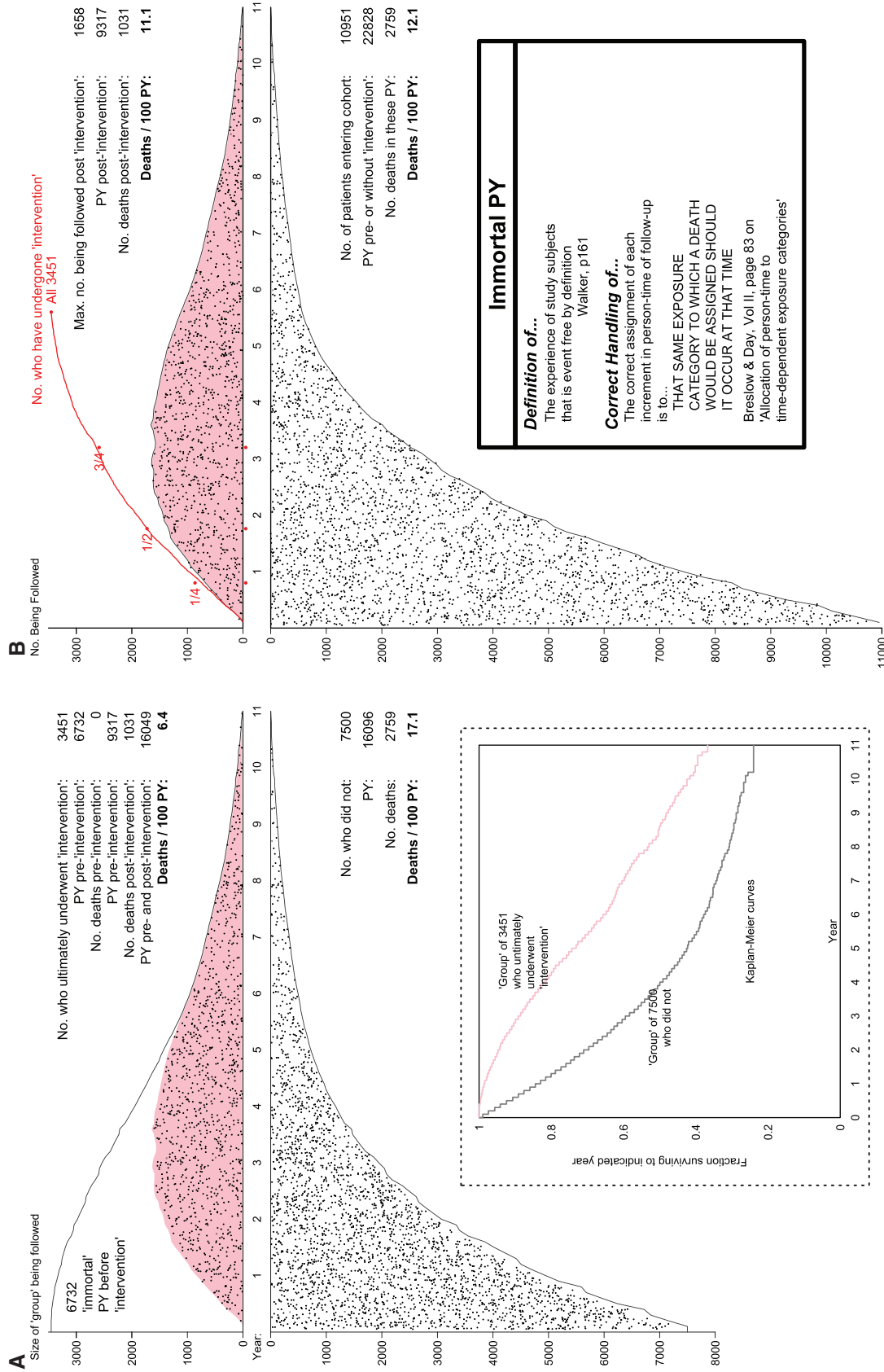


Figure 2. Mortality rates and rate ratios produced by the (A) mis- and (B) proper allocation of pre- and post-intervention patient years. As explained in Figure 1, the hypothetical data for the 10 951 patients were constructed to have an average mortality rate of 3785 deaths in $(10951 \times 2.93 = 32,086)$ patient-years (PY), i.e. 11.8 deaths per 100PY (as in the actual study), but with no variation over years of follow-up, or between states (no, or pre-'intervention' (white background) and post-'intervention' (pink background)). Indeed, the selection of those who changed states (from white to pink polygon, in B) was made at random, and retroactively. The time location (relative to when the allograft failed) of each death is indicated by a black dot. In B, upper panel, the number being followed at any time is smaller than 3451 because some who had received the 'intervention' were already dead before the last ones received it.

long enough to undergo it. One can see how, with some epidemiologists' penchant for long lists of biases, they might term this artefact 'reverse causality bias' (Senn, in a personal communication regarding reference 13, suggested that the 'higher mortality rates' in 'the childless' could equally be reported under the headline: 'Those who die young have fewer children'.)

In this admittedly over-simplified version of the data, with no covariates, the inappropriate analysis led to an apparent rate ratio of $6.4/17.1 = 0.37$. The corresponding 'reduction' of 63%, and an 'improved survival' of at least 2.5 years (areas under the first 11 years of the Kaplan–Meier curves of 7.7 vs 5.2 years), would have been interpreted as having been produced by the intervention, whereas they are merely artefacts of the misallocation of the PY.

Figure 2B shows an appropriate comparison of mortality rates in time-dependent states. With 'each unit of person-time allocated to the state in which the death would have been assigned should it occur at that time', the appropriate rates are (apart from random error) identical, mimicking the theoretical rates used to generate these hypothetical data. The theoretical rates were—unrealistically—taken as constant over the follow-up years. In reality, the PY in each year of follow-up time would be contributed by individuals who were almost one year older than the individuals who contributed PY the year before, and so the mortality rates in successive time-slices would also be successively higher. Thus, since the person-years in the post-'intervention' state are 'older' person-years, a summary rate ratio computed using matched slices of follow-up time would be more appropriate than a crude rate ratio. One would also need to match the person-years on several patient-related factors. As time-slices become more individualized, the distinction between Poisson regression, with its emphasis on the time interval, and Cox's approach, with its focus on the time moment, becomes more blurred. Space does not allow us cover these approaches here, but below (at the end of this article) we provide a link to some additional material we prepared on this topic.

Teachings against such blunders

Warnings against this error go back at least to the 1840s, when William Farr² reminded sanitarians and amateur epidemiologists that:

Certain professions, stations, and ranks are only attained by persons advanced in years; ... hence it requires no great amount of sagacity to perceive that 'the mean age at death', or the age at which the greatest number of deaths occurs, cannot be depended upon in investigating the influence of occupation, rank, and profession upon health and longevity.

Then, in an admirable style seldom equalled in today's writings, he explained that:

If it were found, upon an inquiry into the health of the officers of the army on full pay, that the mean age at death of Cornets, Ensigns, and Second-Lieutenants was 22 years; of Lieutenants 29 years; of Captains 37 years; of Majors 44 years; of Lieutenant-Colonels 48 years; of general Officers, ages still further-advanced ... and that the ages [at death] of Curates, Rectors, and Bishops; of Barristers of seven years' standing, leading Counsel and venerable Judges ... differed to an equal or greater extent ... a strong case may no doubt be made out on behalf of those young, but early-dying Cornets, Curates, and Juvenile Barristers, whose mean age at death was under 30! It would be almost necessary to make them Generals, Bishops, and Judges—for the sake of their health.

Crediting the years of immortality required to reach the rank that the person has reached by the time (s)he dies or follow-up ends exaggerates any longevity-extending benefits of reaching this rank. Likewise, crediting the time until one receives a medical intervention to the intervention exaggerates its life- or time-extending power.

Whereas Farr adopted a tongue-in-cheek style, Bradford Hill³ spelled out the reason for the longevity difference: 'Few men become bishops before they have passed middle life, while curates may die at any age from their twenties upwards'. Separately,⁴ Hill also addressed the fallacy under the heading 'Neglect of the period of exposure to risk':

A further fallacy in the comparison of the experiences of inoculated and uninoculated persons lies in neglect of the time during which the individuals are exposed first in one group and then in the other. Suppose that in the area considered there were on Jan. 1st, 1936, 300 inoculated persons and 1000 uninoculated persons. The number of attacks are observed within these two groups over the calendar year and the annual attack-rates are compared. This is a valid comparison so long as the two groups were subject during the calendar year to no additions or withdrawals. But if, as often occurs in practice, persons are being inoculated during the year of observation, the comparison becomes invalid unless the point of time at which they enter the inoculated group is taken into account.

Hill used a worked example to warn that 'neglect of the durations of exposure to risk must lead to fallacious results and must favour the inoculated'. The example shows that the adjective 'immortal' time is not broad enough: 'event-free time, by definition or by construction' (see Walker, below⁹) is a more general and thus a more appropriate term.

Ten years earlier, Hill⁵ had addressed the ‘period of exposure to risk’ when comparing, ‘from age 25 to age 80’, the longevity of cricketers with that of the general male population.

The comparisons show that cricketers form by no means a short-lived population, but on the contrary hold a substantial advantage at every age... this advantage is undoubtedly somewhat exaggerated since it is assumed that all cricketers are ‘exposed’ from age 25, while in actual fact probably some do not ‘enter exposure’ in first-class cricket till a later age.

Breslow and Day⁶ use a diagram, and a simplified occupational epidemiology example, modelled on the blunder by Duck *et al.*,⁷ to emphasize the correct allocation of person-time, and the distortions produced by misallocation. In Figure 3 we illustrate the Duck *et al.* error and the reallocation of the person-years by Wagoner *et al.*;⁸ our preference for vertical rather than horizontal shading is meant to illustrate the ‘as you go’ (vertical) rather than ‘after the fact’ (horizontal) accumulation of person-years. We also repeat Breslow and Day’s succinct enunciation of the general principle to be followed.

We understand that the term ‘immortal time’ had been used by George Hutchison in the 1970s already, but his Harvard colleague Alexander Walker⁹ is the first we know of to have put the term in writing, in his 1991 textbook. Walker’s numerical examples all involve the correct allocation of such time, with no example given of the consequences of misallocation. The two editions of the Rothman and Greenland textbook¹⁰ do have an example—albeit hypothetical—of the difference between incorrectly and correctly calculated rates based on two parallel groups of exposed and unexposed persons, and state the principle: ‘If a study has a criterion for a minimum amount of time before a subject is eligible to be in the study, the time during which the eligibility criterion is being met should be excluded from the calculation of incidence rates’. They also allude, without an example, to the more general situation where subjects change exposure categories. Unfortunately, it is in this latter situation that most immortal-time blunders are made.

By using the term ‘immortal time’ in the title of a 2003 article, Suissa¹¹ immortalized the term itself. Since then, more than a dozen articles and letters by him and his pharmaco-epidemiology colleagues have addressed the growing number of serious ‘immortal time’ errors in this field. Typically, cohort membership in these studies was defined at the time of diagnosis with, or hospitalization for, a medical condition. The blunders were created by dividing the patients into those who were dispensed a pharmacological agent at some time during follow-up and those who were

not (Unlike in most clinical trials, but like Hill’s inoculation example, not all received it immediately at entry to the cohort.) When, instead, each patient’s follow-up time is correctly divided into the portions where the event-rate of interest might be affected, and the portion where it cannot, the rate-lowering power of the agent disappears.

In several of their articles, Suissa and co-authors use other real datasets to address the same question, and show the consequences of the misallocation. Our annotated bibliography gives several other examples (and collections of examples), by yet other authors, of time-blunders in several other fields. However, even with warnings in one’s own journals, time-blunders continue to occur: 1 year before it received the manuscript containing the study of transplant nephrectomy, the *Journal of the American Society of Nephrology* published an expository article¹² explaining how such a blunder can be recognized and avoided.

Recognizing and avoiding immortal-time blunders

Table 1 lists some ways to recognize immortal time and to avoid the associated traps. We suspect that some of the blunders stem from the tendency—no matter the design—to refer to ‘groups’, as though—in a parallel-arm trial—they were formed at entry and remained closed thereafter. Even when describing a cross-over trial, authors mistakenly refer to the treatment group and the placebo group, rather than to the time when the (same) patients were in the treatment or placebo conditions or states. This tendency may reflect the fact that many questions of prognosis can only be studied experimentally by parallel group designs. Except in studying the short-term effects of alcohol and cellphone use while driving, or medication use or inactivity on blood clots, cross-over designs (called split-plot designs in agriculture) are rare; and their statistical results are more difficult to show graphically and in tables than are those that use independent ‘groups’.

Just as in the story of Solomon, it is appropriate that persons remain indivisible. However, in epidemiology many denominators involve amounts of time (yes, contributed by persons, but time nonetheless), and time is divisible, just as are any other (area- or volume-based) denominators that produce Poisson numerators (The numerators are not divisible.) Despite this, many epidemiologists are less comfortable with dividing an individual’s time into exposed and unexposed portions than they are with measuring research staff size in full-time-equivalents, or than telephone companies are in measuring the amount of time used by customers. We look forward to the companies providing researchers with access to their information on the moment-by-moment location of users’ cellphones,

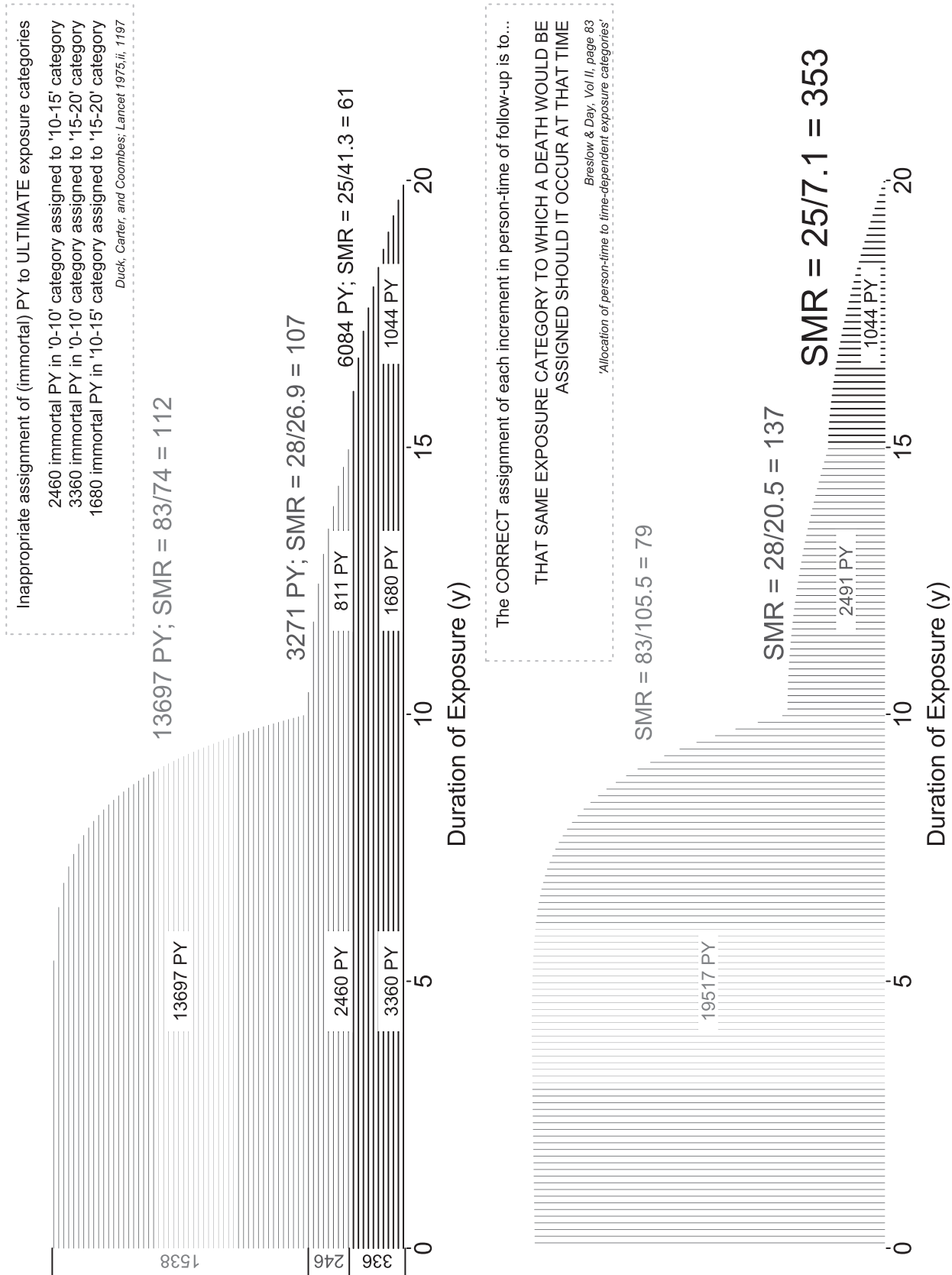


Figure 3. Incorrect (at termination) and correct (as time progresses) allocation of follow-up time in the Duck *et al.* study.⁷ Horizontal timelines represent exposure to vinyl chloride, with durations categorized into 0-10, 10-15 and 15+ years. See also the small worked example, based on this study, in Breslow and Day.⁶

Table 1. Ways to recognize immortal time

Suggestion	Remarks/tests
Distinguish state from trait	A trait (e.g. blood group) is usually forever; people and objects move between states (on/off phone; intoxicated/not; on/off medication; failed allograft in place/removed)
Distinguish dynamic from closed population	Membership in a closed population (cohort) is initiated by an event (transition from a state) and is forever; in a dynamic population, it is for the duration of a state. Dynamic populations are the only option for studying transient exposures with rapid effects (e.g. cellphone/alcohol use vs the rate of motor vehicle accidents)
Focus on person-time in index and reference categories, rather than on people in exposed and unexposed 'groups'	These refer to exposure categories, not to people per se; a person's time may be divided between exposure categories; unless people remain in one category, it is misleading to refer to them as a 'group'
If authors used the term 'group', ask ...	When and how did persons enter a 'group'? Does being in or moving to a group have a time-related requirement? Is the classification a fixed one based on the status at time zero, or later? Is it sufficient to classify a person just once, or do we need to classify the 'person-moments,' that is the person at different times?
Sketch individual timelines	If there are two time scales, a Lexis diagram can help; use different notation for the time portion of the timeline where the event-rate of interest might be affected, and the portion where it cannot (see Figures)
Measure the apparent longevity- or time-extending benefits of inert agents/interventions	After the fact, use a lottery to assign virtual (and never actually delivered) interventions, but with same timing as the one under study. Or use actually-received agents with same timing
Imagine this agent/intervention were being tested within a randomized trial	How, and when after entry, would the agent be assigned? Administered? How would event rates be computed? How would Farr have tested his 'early-promotion' suggestion?
Think short intervals and hazard rates, even if the hazard rates do not change abruptly	In addressing the present, conditional on the past, the hazard approach has already correctly documented the experience in each small past interval; the natural left to right time-ordering of the short intervals allows for correct recognition of transitions between exposure states. By computing a mortality rate over a longer time-span defined after the fact, one may forget that in order to contribute time to the index category, people had to survive the period spent in the (initial) reference category

so that they can more accurately measure the amounts of on-the-phone and off-the-the phone driving time, and the rates of motor vehicle accidents in these. The more comfortable biomedical researchers become in dividing an individual's time (e.g. same person with different hearts), the less the risk of immortal time blunders.

It is our impression that epidemiologists are more comfortable 'splitting time' than many researchers in the social sciences, where correlations (rather than differences in and ratios of incidence rates) are the norm. If one is studying the duration of life, using lives that have been completed, it may not matter much whether one compares the aggregated lives divided by their number (average lifetime) or the total number (of deaths) divided by the total lifetime (the average number of deaths per unit of time). However, once one restricts attention to the rate of (terminating) events within just portions of these lifetimes, the switching between 'exposure' states, incomplete lives and censored and truncated observations all make it much more difficult to stay with the familiar correlations carried out using the time scale itself. The 'correlations between election age and death age for restricted subsamples based on election age

percentile'¹³ and the 'setting time-zero' to some arbitrary birthday (e.g. 0, 50 or 65 in the case of those nominated for an Oscar) are good examples of the limitations of staying with the average duration (longevity) scale that is easier to convey to the public. Those who compare rates (dimension: time⁻¹) within the relevant time-windows have much more flexibility than those who attempt to compare average durations (dimension: time).

Theories such as the just-cited precocity-longevity hypothesis are seductive, and have a certain plausibility. But some of this may be a result of the framing. A restatement of the 'evidence' can help uncover the fallacy: imagine if Groucho Marx were to re-word it, using Ronald Reagan's election and longevity as the example. In any case (as we stated in our re-examination of the claimed 3.9 year longevity advantage for Oscar winners see additional references), no matter how important or unimportant results would be if they were true, 'readers and commentators should be doubly cautious whenever they encounter statistical results that seem too extreme to be true'.

Failure to recognize immortal time errors leads to consequences that in some cases may be serious and costly,

and not easily corrected. In a case like the pharmaco-epidemiology work of Suissa and colleagues, the costs of correcting the record can be considerable. Original stories in the lay press are not usually updated: the Oscar longevity story in the *Harvard Health Letter* of March 2006 is still available online—if one is willing to pay \$5 for access. The Harvard website does not cite any updates, nor does the (itself widely cited) 2010 *New York Times* interview.¹⁴ *Forbes Magazine*¹⁵ is an exception. Medical journal editors, too, appear reluctant to correct blunders missed by peer review. Indeed, when one of us (B.F.) wrote to the Editor of the *American Journal of Nephrology* to point out the strong possibility that the finding of ‘improved survival’ following allograft nephrectomy was an artefact, she was told that the journal did not have a Letters to the Editor section, but that it would pass on the concerns to the authors.

Given HR Haldeman’s observation, ‘Once the toothpaste is out of the tube, it’s hard to get it back in’, it seems prudent and scientifically responsible to try to avoid immortal time errors from the outset. Researchers can avoid ‘immortal time’ errors by classifying person-time into exposure states, rather than classifying whole persons who ultimately attain an exposure state into ‘exposed’ and ‘unexposed’ groups, under the assumption that they have been in those groups from the outset.

Or, as Steve Jobs told us, ‘Think different’. Think person-time, not person.

Additional material

Since the ‘extended’ Cox model is often used in this ‘change of states’ context, our first version of this manuscript contained a section entitled ‘Data analysis options’, illustrated with ‘survival times after cardiac allografts’, taken from a classic article on survival post heart-transplant. That section, and the associated computer code, can be found on the author’s website <http://www.medicine.mcgill.ca/epidemiology/hanley/software>.

In that material we show how we would deal with these data today, using time-dependent covariates in a multivariable parametric or semi-parametric (hazard) regression model, with subjects switching ‘exposure’ categories (states) over time. However, we found it instructive to begin with the classical approaches already widespread in 1969, in particular those in Mantel’s classic 1959 article. We use his 1974 generalization of lifetables¹⁸ to deal with transitions between ‘exposure’ states; indeed, Mantel’s 1974 paper is the conceptual forerunner of what is now known as regression for ‘time-varying covariates’. We also estimate the mortality rates and rate ratios using Poisson regression.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada, and by the Canadian Institutes of Health Research.

Conflict of interest: None declared.

Postscript

When writing this piece, we wondered whether we were preaching to the converted. We did not add Rodolfo Saracci’s suggested subtitle, ‘The fallacy that refuses to die’. Surely such blunders do not occur in epidemiology journals, where the review is more rigorous than in some of the clinical ones? The article that is the subject of the correspondence in this IJE issue¹⁹ indicates otherwise. The flaw in the comparison that led to a multifactorially adjusted, but too good to be true, hazard ratio of 0.52 (and even the other, more finely stratified ratios) was missed not just by the authors themselves, but also by their colleagues, granting agencies, journal referees and editors, and newspaper journalists and editors.

The Editor asked us to ‘explain how immortal time bias plays a role in their findings’ and to provide ‘any comment [we] care to make about their re-analysis in response²⁰ to the criticisms raised by Lange and Keiding’.²¹ We do so, but only after we first make some broader comments.

It will not be easy to put the toothpaste back in the tube, but we hope that those in the academic portion of this chain will each do their part. Might the *IJE* ask its media contacts to carry a follow-up story that might help undo the damage? In addition, instead of reporting additional analyses that still have flaws (or faulting the media for the over-interpretation and for their focus on the longevity ‘effect’) an *IJE mea (nostra?) culpa* might do more good: it might just add to (rather than subtract from) the limited amount of credibility biomedical scientists currently have remaining with the public.

It is one thing to give the public a reason to merely day-dream about winning an Oscar and adding four years to one’s life; it is quite another to imply—even cautiously—on the basis of the difference in median longevity of six years in the bottom left panel of Figure 1 of the ‘sun exposure’ article, that an even larger longevity bonus is readily accessible to all. Curiously, the ‘extra’ six years do not appear anywhere in the article, but figured prominently in the newspaper story. In it, one of the authors emphasized that they could not identify the direct causal link, but added that ‘the numbers as such do not lie’. This statement illustrates what one might call a type III error, where an inappropriately set up statistical contrast, not chance, is the culprit.

We comment later on the less-emphasized, but possibly valid, results in Table 3 of the article, and first address the contrasts that led to the crude difference of six years and the ‘adjusted’ hazard ratio of 0.52. How could these analyses have received the most prominence, and without anyone ever raising an alarm? There was a hint that the authors understood, on some level, that such analyses involved immortal time: there is a statement about ‘the temporality between the exposure and the outcomes’. The use of logistic regression for two of the outcomes, but Cox regression for the other, should also have prompted reviewers to try to understand why. In retrospect, the warning signs were all there: *P*-values so small that—even setting aside the concern already raised, tongue in cheek, about their numerical accuracy—they may be the smallest that have ever appeared in print anywhere; very different answers from the various data-analysis approaches; Kaplan–Meier curves and log-rank tests with no mention of staggered entry, but Cox regressions that do; the quite telling pattern of hazard ratios in the lower left panel of Figure 2; and, most importantly, at least to those not involved in the publication chain, the six years and even the adjusted hazard ratio of 0.52 seemed way too good to be true (in industrialized countries the mortality rate ratio (females:males) is approximately 0.7). On the other hand, as with the possibility of burnout of presidents who are early achievers, or of the extra health benefits of being rich, or of taking out a failed transplant, the ‘more-(activity-in-the-)sun-is-good’ hypothesis has a certain plausibility to it, and there is other evidence, based on the 10-year survival of those with basal cell carcinoma (Jensen *et al.*).²² Moreover, the authors had used a clever (if somewhat unusual) way to study it, and used sophisticated statistical tools with extensive high quality data. The consistency in those below age 90 (‘less MI, less fracture, less death’) was taken as further evidence in support of the hypothesis.

One way to ‘check’ for immortal time bias is to study an event (outcome) that should have no causal relationship with the exposure of interest, and to be wary if the hazard ratios are clearly below 1. Alternatively, one may examine the association between a clearly ‘unrelated’ exposure and the outcome of interest, as we do below.

The article has conflicting descriptions of what was done to generate the less-emphasized results in Table 3. The statistical methods describe (synchronized?) matched sets, each comprising one ‘exposed’ person and five persons referred to as ‘general population controls’ but in the results section it is referred to as a ‘matched case-control study’. Importantly, the authors do tell us that ‘only myocardial infarction and hip fracture events following a diagnosis of non-melanoma skin cancer or cutaneous malignant melanoma entered into the analysis, whereas

events before skin cancer were excluded’. This, together with the (adjusted) all-cause mortality hazard ratios of 0.96 and 0.97, suggest that, whatever they called it, their analysis may have—partially at least—avoided the ‘temporality’ problem (As we illustrate below, the crude six years, and the adjusted 0.52, and even the hazard ratios in the authors’ additional analyses do not). The authors now realize that the analyses in Table 3, initially relegated to the very end of the Results and not discussed further, are probably the least biased. If one takes the fully adjusted 0.97 from the matched study as the closest to correct, one could put it into context for the newspaper readers by saying that it translates into a longevity difference of about four months rather than several years.

To show how immortal time bias plays a role in their findings, and to try to understand if the additional analyses are free of it, we examine the association between an unrelated exposure and death from any cause. Retroactively, and randomly, and without communicating the information to anyone, we choose a number of anonymous Danes in the Lexis rectangle enclosed by ages 40–110 years and calendar years 1980–2006 to be ‘prizewinners’; winning the prize is the new, and obviously ‘irrelevant’ exposure. The population size Lexis dataset available in the Human Mortality Database (<http://www.mortality.org>) has a total of 4 130 227 Danes (the survivors, past age 40 and past 1980, of 91 different birth cohorts) in the leftmost column or bottom row of the rectangle. Using R code (available on our website) we simulated a yearly lottery that selected some of them to be virtual prizewinners. The incidence of prizewinners was an age-function with the same shape as the age-specific incidence of non-melanoma skin cancer in several Canadian provinces, scaled (downwards!) so that the total number of winners, and the average age of winning, were close to the 129 000 cases of skin-cancer, and the average diagnosis age of 68, in the *IJE* article. The only condition was that the winner had to be alive at the time of each yearly draw. Unlike other lotteries, there was, in large print, a statement that ‘no other conditions apply’.

By its nature, the prize could not extend their longevity. Yet, just because of this ‘must be living’ condition, when we used the same analysis as in Figure 1 in the *IJE* article, we obtained a difference in median longevity of 8.5 years (and a hazard ratio of 0.57 with a *P*-value somewhere below the R `pchisq` function limit of 5×10^{-324}). Moreover, the hazard ratios we found in the 10-year ‘strata’ looked very similar to those in the lower left panel in the *IJE* Figure 2. Furthermore, when (as the authors do in their response) we narrowed the age slices further and insisted that ‘those who [won our prize] beyond the age-strata enter into the analysis as not having [won]’, we again get patterns similar to those in the figure in the

response to Lange and Keiding. Even using age-slices just two years wide, our hazard ratios were not null: they ranged from 0.93 at age 65 to 0.95 at age 85. The reason for the residual bias is that, by definition, a person who receives the prize at age 77.9 is 'immortal' for 1.9 years of the 2-year age slice 76–78. To avoid this induced immortality entirely, one needs to shrink the age-slice to an instant. Doing so is equivalent to using a time-dependent covariate ('exposure') in the Cox model, with risk sets defined at the moments the events occur. This is the most common way to deal with exposure states rather than traits.

By matching the Cox models on age, the authors did compare people who have survived to the same age. This may have led them and the reviewers to think that all was now taken care of. But with changing exposures, age-matching alone is not sufficient: one must also properly identify and update each subject's unexposure status as he/she proceeds through time and through the risk sets.

Sadly, we must add one more example to the list begun by Farr: to enter the index category, i.e. be promoted in one's profession; enter the list of cricket or jazz greats; enter the period of exposure to risk; receive an organ transplant; have it removed; be prescribed inhaled corticosteroids or a statin; win an Oscar; or receive a diagnosis of non-melanoma skin cancer, one needs to have lived long enough (in the reference category) in order to do so. No such minimum longevity requirement is imposed on entry to the reference category itself.

Annotated references

1. Ayus JC, Achinger SG, Lee S, Sayegh MH, Go AS. Transplant nephrectomy improves survival following a failed renal allograft. *J Am Soc Nephrol* 2010;21:374–80. See also the editorial in the same issue. We provide extensive commentary in sections 2 of our article, and in the Supplementary material.
2. Farr W. *Vital Statistics*. A memorial volume of selections from the reports and writings of William Farr. London: The Sanitary Institute, 1885.
The now-easy access, clear writing, and opportunities to compare the concepts and principles with those in modern textbooks, make several portions of this volume worth studying even today.
3. Hill AB. Principles of medical statistics. XIV: Further fallacies and difficulties. *Lancet* 1937;229:825–27.
'The average age at death is not often a particularly useful measure. Between one occupational group and another it may be grossly misleading ... the average age at death in an occupation must, of course, depend in part upon the age of entry to that occupation and the age of exit from it—if exit takes place for other reasons than death'.
4. Hill AB. Principles of medical statistics, XII: Common fallacies and difficulties. *Lancet* 1937;229:706–08.
With the dates changed, the worked example could easily pass for a modern one: 'Suppose on Jan. 1st, 1936, there are 5000

persons under observation, none of whom are inoculated; that 300 are inoculated on April 1st, a further 600 on July 1st, and another 100 on Oct. 1st. At the end of the year there are, therefore, 1000 inoculated persons and 4000 still uninoculated. During the year there were registered 110 attacks amongst the inoculated persons and 890 amongst the uninoculated, a result apparently very favourable to inoculation. This result, however, must be reached even if inoculation is completely valueless, for no account has been taken of the unequal lengths of time over which the two groups were exposed. None of the 1000 persons in the inoculated group were exposed to risk for the whole of the year but only for some fraction of it; for a proportion of the year they belong to the uninoculated group and must be counted in that group for an appropriate length of time.'

A mathematical proof that 'neglect of the durations of exposure to risk must lead to fallacious results and must favour the inoculated' can be found in the paper by Beyersmann *et al.*, *J Clin Epidemiol* 2008;61:1216–21. Hill goes on to describe a 'cruder neglect of the time-factor [that] sometimes appears in print, and may be illustrated as follows. In 1930 a new form of treatment is introduced and applied to patients seen between 1930 and 1935. The proportion of patients still alive at the end of 1935 is calculated. This figure is compared with the proportion of patients still alive at the end of 1935 who were treated in 1925–29, prior to the introduction of the new treatment. Such a comparison is, of course, inadmissible'. Today's readers are encouraged to compare their reason why with that given by Hill.

5. Hill AB. Cricket and its relation to the duration of life. *Lancet* 1927;949–950.
6. Breslow NE, Day NE. *Statistical Methods in Cancer Research: Volume II: The Design and Analysis of Cohort Studies*. New York: Oxford University Press, 1994.
The basis for Figure 3, and the source of the principle by which person-time should be allocated.
7. Duck BW, Carter JT, Coombes EJ. Mortality study of workers in a polyvinyl-chloride production plant. *Lancet* 1975;2:1197–99. 'Age-standardised mortality-rates for a population of 2100 male workers exposed to vinyl chloride for periods of up to 27 years do not show any excess of total or cause-specific mortality'.
8. Wagoner JK, Infante PF, Saracci R. Vinyl chloride and mortality? [Letter] *Lancet* 1976;2:194–95.
This letter, a response to Duck *et al.*, is notable both for its tongue-in-cheek 'possible interpretations' of the SMRs of 112, 107 and 61 (!) in the PY where exposure had been for <10, 10–15 and 15+ years, respectively, and for its re-allocation of the PY giving new SMRs of 79, 137 and 353!
9. Walker AM. *Observation and Inference: An Introduction to the Methods of Epidemiology*. Chestnut Hill, MA: Epidemiology Resources, 1991.
This is the first author we know of to define and put the term 'immortal time' in writing. Readers are invited to compare the definition with the description provided by Bradford Hill.
10. Rothman KJ, Greenland S. *Modern Epidemiology*. 2nd edn. Philadelphia, PA: Lippincott-Raven 1998.
11. Suissa S. Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. *Am J Respir Crit Care Med* 2003;168:49–53.

- The first author to put the term 'immortal time' in the title of an article.
12. Shariff SZ, Cuerden, MS, Jain AK, Garg AX. The secret of immortal time bias in epidemiologic studies. *J Am Soc Nephrol* 2008;19:841–43.
This article appeared one year before, in the same journal as the one on transplant nephrectomy.
 13. McCann SJH. The precocity-longevity hypothesis: earlier peaks in career achievement predict shorter lives. *Pers Soc Psychol Bull* 2001;27:1429–39.
See section 4, and see two subsequent articles by McCann. Does the fact that President Ronald Reagan lived to be 93 support this hypothesis? 'Likewise, Stephen Senn (personal communication) suggests that the 'higher mortality rates' in 'the childless' (Agerbo *et al.* Childlessness, parental mortality and psychiatric illness: a natural experiment based on in vitro fertility treatment and adoption. *J Epidemiol Community Health* 2012;00:1–3), could equally be reported under the headline 'Those who die young have fewer children'.
 14. <http://www.nytimes.com/2010/08/31/science/31profile.html> see also http://test.causeweb.org/wiki/chance/index.php/Oscar_winners_do_not_live_longer
 15. Matthew Herper. Does winning an Oscar make you live longer? Three important lessons from a raging scientific debate. *Forbes Magazine*. Feb 26, 2012.
 16. Leibovici L. Effects of remote, retroactive intercessory prayer on outcomes in patients with bloodstream infection: randomised controlled trial. *BMJ* 2001;323:1450–51.
We modelled our simulation on this (*BMJ* Christmas Edition) article, which generated many intense responses. It began with the statement 'As we cannot assume a priori that time is linear, as we perceive it, or that God is limited by a linear time, as we are, the intervention was carried out 4–10 years after the patients' infection and hospitalisation'.
 17. Turnbull BW, Brown BW, Hu M. Survivorship analysis of heart-transplant data. *J Am Stat Assoc* 1974;69:74–80.
Their permutation test motivated us to simulate, by lottery, a virtual, after-the-fact and ineffective intervention for those with a failed allograft. It is a useful technique for quantifying how much of the longevity advantage is an artefact. See our earlier use of this strategy in Sylvestre M-P, Huszti E, Hanley JA. Do Oscar winners live longer than less successful peers? A reanalysis of the evidence. *Annals of Internal Medicine* 2006; 145:361–63.
 18. Mantel N, Byar DP. Evaluation of response-time data involving transient states illustration using heart-transplant data. *Journal of the American Statistical Association* 1974;69:81–86.
 19. Jacobsen P, Nordestgaard B, Nielsen S, Benn M. Skin cancer as a marker of sun exposure associates with myocardial infarction, hip fracture, and death from any cause. *Int J Epidemiol* 2013;42:1486–96.
 20. Brøndum-Jacobsen P, Nordestgaard BG, Nielsen SF, Benn M. Authors' Response to: Skin cancer as a marker of sun exposure—a case of serious immortality bias. *Int J Epidemiol* 2014;43:972–73.
 21. Lange T, Kielsing N. Letter: Skin cancer as a marker of sun exposure: a case of serious immortality bias. *Int J Epidemiol* 2014;43:971.
 22. Jensen AØ, Lamberg AL, Jacobsen JB, Braae Olesen A, Sørensen HT. Non-melanoma skin cancer and ten-year all-cause mortality: a population-based cohort study. *Acta Derm Venereol* 2010;90:362–67

Additional references

Redelmeier DA, Singh SM. Survival in Academy Award-winning actors and actresses. *Ann Intern Med* 2001;134:955–62.

The widely reported almost four-year longevity advantage over their 'nominated but never won' peers includes the immortal years between being nominated and winning. Use of the years between birth and nomination is an example of many researchers' reluctance to subdivide each person's relevant experience. See the re-analyses by Sylvestre *et al.* (2006) in the same journal, and by Wolkewitz *et al.* (*Am Statistician* 2010;64:205–11) and Han *et al.* (*Applied Statistics* 2011;5:746–72).

Mantel N, Byar DP. Evaluation of response-time data involving transient states—illustration using heart-transplant data. *J Am Stat Assoc* 1974;69:81–86.

In 1972, Gail had identified several biases in the first reports from Houston and Stanford. One was the fact that 'patients in the [transplanted] group are guaranteed (by definition) to have survived at least until a donor was available, and this grace period has been implicitly added into [their] survival time'. Mantel was one of the first to suggest statistical methodologies for avoiding what he termed this 'time-to-treatment' bias, where 'the survival of treated patients is compared with that of untreated controls, results from a failure to make allowance for the fact that the treated patients must have at least survived from time of diagnosis to time of treatment, while no such requirement obtained for their untreated controls'. He introduced the idea of crossing over from one life table ('waiting for a transplant' state) to another ('post-transplant') and make comparisons matched on day since entering the waitlist. Incidentally, Mantel's choice of the word 'guarantee' is not arbitrary: textbooks on survival data refer to a 'guarantee time' such that the event of interest many not occur until a threshold time is attained. In oncology trials, a common error—usually referred to as 'time-to-response' or 'guarantee-time' bias—is to attribute the longer survival of 'responders' than 'nonresponders' entirely to the therapy, and to ignore the fact that, by definition, responders have to live long enough for a response to be noted (see Anderson, *J Clin Oncol*. 1983;1:710–19, and, more recently, Giobbie-Hurder *et al.* *J Clin Oncol* 2013;31:2963–69).

Glesby MJ, Hoover DR. Survivor treatment selection bias in observational studies: examples from the AIDS literature. *Ann Intern Med* 1996;124:999–1005.

'Patients who live longer have more opportunities to select treatment; those who die earlier may be untreated by default' and their three words 'survivor treatment selection', to describe the bias explain why some person-time is 'immortal'.

Wolkewitz M, Allignol A, Harbarth S, de Angelis G, Schumacher M, Beyersmann J. Time-dependent study entries

and exposures in cohort studies can easily be sources of different and avoidable types of bias. *J Clin Epidemiol* 2012;65:1171–80.

Using an example from hospital epidemiology, the authors give ‘innovative and easy-to-understand graphical presentations of how these biases corrupt the analyses via distorted time-at-risk’. See also: Schumacher *et al.* Hospital-acquired infections—appropriate statistical treatment is urgently needed! *Int J Epidemiol* 2013;42:1502–08.

Rothman KJ. Longevity of jazz musicians: flawed analysis. [Letter]. *Am J Public Health* 1992;82:761.

A letter in response to a retired professor of management, and jazz amateur (but sadly also a statistical amateur), whose data analysis suggested that jazz musicians, despite their rough lifestyle, live at least as long as their peers. In ‘Premature death in jazz musicians: fact or fiction?’ (Spencer FJ. *Am J Public Health* 1991;81:804–05), the longevity of their peers was measured by the life expectancy of those born the same year as they, although the musicians are, by definition, immortal until they became musicians and eminent enough to be included in the sample. The tone of the letter provides also an interesting contrast with Farr’s teaching style. See: Bellis *et al.* Elvis to Eminem: quantifying the price of fame through early mortality of European and North American rock and pop stars. *J Epidemiol Community Health* 2007;61:896–901; Hanley *et al.* How long did their hearts go on? A Titanic study. *BMJ* 2003;327:1457; Abel *et al.* The longevity of Baseball Hall of Famers compared to other players. *Death Studies* 2005;29:959–63.; Redelmeier *et al.* Death rates of medical school class presidents. *Soc Sci Med* 2004;58:2537–43; and Olshansky SJ. Aging of US Presidents. *JAMA* 2011;306:2328–29, for more appropriate ways to carry out such longevity comparisons.

van Walraven C, Davis D, Forster AJ, Wells GA. Time-dependent bias was common in survival analyses published in leading clinical journals. *J Clin Epidemiol* 2004;57:672–82.

They gave immortal time bias a slightly different name because they covered a slightly broader spectrum of situations. Their review surveyed articles containing survival analysis that may have incorrectly handled what they define as a ‘baseline immeasurable’ time-dependent variable, i.e. one that could not be measured at baseline. They focused not just on the exposure of interest, but also other time-dependent covariates. They describe

an interesting study on whether patients having a follow-up visit with a physician who had received the discharge summary would have a lower rate of re-hospitalization. When analysed as a fixed-in-time variable (i.e. as two ‘groups’, we found a large difference in readmission rates. However, this is a biased association, because patients who are readmitted to the hospital early after discharge do not have a chance to see such physicians and are placed in the ‘no-summary’ group. When a (correct) time-dependent analysis is used, we found a much smaller rate difference. They examined a large number of observational studies that used a survival analysis, including the one on the survival of Oscar nominees and winners. The only ‘baseline immeasurable’ time-dependent covariate in that study the reviewer(s) identified was whether actors changed their names after baseline; missed was the fact that some nominees who did not win the first time they were nominated changed to the winners category subsequently. So, they ‘cleared’ this article, declaring it to be free of any possible time-dependent bias. Interestingly, they also thanked one of the authors of the Oscar study for ‘comments regarding previous versions of this study’.

Carrieri MP, Serraino D. Longevity of popes and artists between the 13th and the 19th century. *Int J Epidemiol* 2005;34:1435–36.

Compared with the bishops in Farr’s example, popes must have survived even longer just to become pope. Even though the authors were aware that longevity is a ‘necessary condition for being elected Pope’, their statistical approach did not fully address this constraint. Ideally, for each papacy-specific ‘longevity competition’, the time clock starts when the pope is elected, and the competition should include the pope, and those artists born the same year as him, who were still alive when he was elected. However, for several papacies, such detailed matching is not possible. Instead, for each of the 1200–1599 papacies, their analysis effectively ‘started the clock’ at age 39—the age at which the youngest pope in that era was elected—by excluding artists who died before reaching that age. For the 1600–1900 papacies, it was started at age 38. A re-analysis (Hanley JA, Carrieri MP, Serraino D. Statistical fallibility and the longevity of popes: William Farr meets Wilhelm Lexis. *Int J Epidemiol* 2006;35:802–05), that used a papacy-specific time clock for each papacy-specific longevity competition, reversed the original findings.

03-Apr-2014

IJE-2012-12-1250.R2 - Avoiding blunders involving 'immortal time'

Dear Drs Hanley and Foster

Many thanks for your excellent article on immortal time bias and further to my note of acceptance yesterday. We recently published an article on sun exposure and longevity (1) (attached) which was subject to this problem and resulted in high levels of publicity.

A critical letter was received followed by a response by the authors (both attached) which we plan to publish in the same issue of IJE in which your article will appear.

Would you be interested in adding some extra text to your paper by way of comment or explanation as to how immortal time bias (or as you prefer 'period of exposure to risk') plays a role in their findings and any comment you care to make about their re-analysis in response to the criticisms raised by Theis Lange and Niels Keidin? We feel this would greatly help our readers in understanding the concept and demonstrating how it is still a rather tricky issue for skilled investigators to deal with.

1. Brondum-Jacobsen P, Nordestgaard BG, Nielsen SF, Benn M 2013 Skin cancer as a marker of sun exposure associates with myocardial infarction, hip fracture and death from any cause. Int J Epidemiol 42:1486-149

Thank you for your support.

All good wishes
Shah Ebrahim
Editor in Chief
International Journal of Epidemiology

Skin cancer as a marker of sun exposure associates with myocardial infarction, hip fracture and death from any cause

Peter Brøndum-Jacobsen,^{1,3} Børge G Nordestgaard,^{1,3} Sune F Nielsen¹ and Marianne Benn^{2,3*}

¹Department of Clinical Biochemistry, Herlev Hospital, Herlev, Denmark, ²Department of Clinical Biochemistry, Gentofte Hospital, Copenhagen, Denmark and ³Copenhagen University Hospital and Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

*Corresponding author. Department of Clinical Biochemistry, Gentofte Hospital, Niels Andersenvej 65, DK-2900 Hellerup, Copenhagen, Denmark. E-mail: Marianne.Benn@regionh.dk

Accepted 22 July 2013

Background Sun exposure is the single most important risk factor for skin cancer, but sun exposure may also have beneficial effects on health. We tested the hypothesis that individuals with skin cancer (non-melanoma skin cancer and cutaneous malignant melanoma) have less myocardial infarction, hip fracture and death from any cause, compared with general population controls.

Methods We examined the entire Danish population above age 40 years from 1980 through 2006, comprising 4.4 million individuals. Diagnoses of non-melanoma skin cancer ($n = 129\,206$), cutaneous malignant melanoma ($n = 22\,107$), myocardial infarction ($n = 327\,856$), hip fracture ($n = 129\,419$), and deaths from any cause ($n = 1\,629\,519$) were drawn from national registries.

Results In individuals with vs without non-melanoma skin cancer, multivariably adjusted odds ratios were 0.96 (95% confidence interval: 0.94–0.98) for myocardial infarction and 1.15 (1.12–1.18) for hip fracture, and the multivariably adjusted hazard ratio was 0.52 (0.52–0.53) for death from any cause. Risk of hip fracture was reduced (odds ratios were below 1.0) in individuals below age 90 years. In individuals with vs without cutaneous malignant melanoma, corresponding odds ratios were 0.79 (0.74–0.84) for myocardial infarction and 0.84 (0.76–0.93) for hip fracture, and the corresponding hazard ratio for death from any cause was 0.89 (0.87–0.91); however, cutaneous malignant melanoma was associated positively with death from any cause in some individuals.

Conclusions In this nationwide study, having a diagnosis of skin cancer was associated with less myocardial infarction, less hip fracture in those below age 90 years and less death from any cause. Causal conclusions cannot be made from our data. A beneficial effect of sun exposure *per se* needs to be examined in other studies.

Keywords Sun exposure, skin cancer, myocardial infarction, hip fracture, mortality, nationwide study

Introduction

Public health recommendations warn against high sun exposure in view of the risk of skin cancer in general and cutaneous malignant melanoma in particular. However, sun exposure has been reported to be associated with lower risk of cardiovascular diseases and with other beneficial effects on health.^{1,2} Although the balance between positive and negative effects of sun exposure in the public debate currently leans towards the negative side, the scientific evidence for this balance is largely unclear.

Sun exposure is the single most important risk factor in the pathogenesis of skin cancer, accounting for an estimated 80–85% of both non-melanoma basal cell carcinoma and squamous cell carcinoma (here collectively referred to as non-melanoma skin cancer) and cutaneous malignant melanoma.^{3,4} Constant and prolonged sun exposure patterns cause non-melanoma skin cancer, whereas overexposure as a child and high intensity intermittent sun exposure primarily cause cutaneous malignant melanoma.^{5,6}

We tested the hypothesis that having a diagnosis of skin cancer was associated with less myocardial infarction, hip fracture and death from any cause, compared with general population controls. We chose to include these three hard outcomes because myocardial infarction and hip fracture (as a clinical marker of osteoporosis) almost always lead to hospitalization in Denmark and therefore are registered as described below, and because these two diagnoses are unlikely to be given to patients during a hospitalization without proper diagnostic tests. Furthermore, death is the hardest of all outcomes and is registered 100% in Denmark. We studied the entire Danish population above age 40 years from 1980 through 2006 and used information from the national Danish Cancer Registry, the national Danish Patient Registry, the national Danish Causes of Death Registry, the national Danish Civil Registration System and Statistics Denmark; all registries were complete during this period. We first used a cross-sectional design for the outcomes myocardial infarction and hip fracture and a prospective design for the outcome death from any cause, and secondly, a matched design to circumvent effects of time (calendar year) and changes in sun exposure habits and in treatment of cancer during the observation period.

Methods

We conducted a study of the entire Danish population above age 40 years from 1 January 1980 through 31 December 2006, comprising 4 412 568 individuals. Almost 90% of the Danish population are Whites of Danish descent. Denmark is situated in the northern hemisphere at latitudes 54–57N and has a mean of 1495 sun-h per year or a mean of 4.1 sun-h per day (www.dmi.dk). The national Danish Civil Registration

System records all births, deaths, emigrations and immigrations in Denmark, recorded by a civil registration number unique to every person living in Denmark, including information about age and gender.

This study was approved by Herlev Hospital, Copenhagen University Hospital, Statistics Denmark and the Danish Data Protection Agency. Anonymous nationwide studies in Denmark do not require approval from ethical committees.

Exposures: non-melanoma skin cancer and cutaneous malignant melanoma

Diagnoses and dates of skin cancer were drawn from the national Danish Cancer Registry, which identifies 98% of cancer cases in Denmark from all hospitals and private practising pathologists; neither non-melanoma skin cancer nor cutaneous malignant melanoma diagnoses were based on self-reports.⁷ All individuals with a diagnosis of non-melanoma skin cancer according to the International Classification of Diseases (ICD-7 until 31 December 2003, thereafter ICD-10; ICD-7: 191; ICD-10: C44) and cutaneous malignant melanoma (ICD-7: 190; ICD-10: C43) from 1 January 1980 through 31 December 2006 were identified.

Outcomes: acute myocardial infarction, hip fracture and death from any cause

Diagnoses and dates of myocardial infarction and hip fracture were drawn from the national Danish Patient Registry and the national Danish Causes of Death Registry, recording information on discharge diagnoses from all Danish hospitals including outpatients and causes of death reported by hospitals and general practitioners using the civil registration number.⁷ Myocardial infarction (ICD-8: 410; ICD-10: I21) and hip fracture (ICD-8: 820; ICD-10: S72.0, S72.1, S72.2) from 1980 through 2006 were used in the study.

Information on death from any cause was drawn from the national Danish Civil Registration System, recording information about deaths in Denmark, using the civil registration number.

Other covariates

Statistics Denmark records information on descent coded as Danish or other descent, educational level and geographical residential city size for all persons living in Denmark. From 1 January 1980 through 31 December 1995, Statistics Denmark also recorded detailed information on occupation with 202 different categories. Each occupational category was assigned an estimated sun exposure level (low or high) and an estimated physical activity level (low, intermediate or high) based on general knowledge, and two variables were generated. For example, farmers will be coded as high occupational sun exposure and high occupational physical activity and office workers will be low in both categories.

Statistical analysis

Statistical analyses were performed with STATA MP 11.1 software. We assessed the association between diagnoses of non-melanoma skin cancer and cutaneous malignant melanoma and the three outcomes, myocardial infarction, hip fracture and death from any cause, by surveillance of all individuals above age 40 years living in Denmark from 1 January 1980, from the 40th birthday or from time of immigration (whichever occurred last) to occurrence of the outcome investigated (e.g. myocardial infarction, hip fracture or death from any cause), emigration or 31 December 2006 (whichever occurred first). Individuals who first emigrated and later returned to Denmark were still included in the analyses. We used Kaplan–Meier curves and log rank tests. For the outcomes myocardial infarction and hip fracture, we used logistic regression models because of the temporality between the exposure and the outcomes, and odds ratios were calculated as measures of relative risk. For the endpoint death from any cause, we used Cox regression models with age as the time scale, implying that age is automatically adjusted for, and hazard ratios were calculated as measures of relative risk. The Cox regression models were left truncated (in 1980, at the 40th birthday or at immigration) with delayed entry, and individuals were censored at event, death, permanent emigration or end of follow-up. We assessed the assumption of proportional hazards graphically by plotting log (cumulative hazards) as a function of follow-up time. We detected no major violations until age 100 years for myocardial infarction, hip fracture or death from any cause, except for cutaneous malignant melanoma and death from any cause. To address potential modification by age, we also performed the above-mentioned analyses in age-strata of 10 years.

Both regression models were adjusted multifactorially for age, gender, descent, geographical residency, educational level, estimated occupational sun exposure and estimated occupational physical activity, and were also stratified by baseline characteristics.

To circumvent the effect of time (calendar year) and changes in sun exposure habits and in treatment of cancer during the past three decades, we performed a matched analysis matching each individual with non-melanoma skin cancer or cutaneous malignant melanoma to five general population controls on the basis of age, birth year and gender; we then used logistic regression modeling overall and in age-strata of 10 years.

Results

We included the entire Danish population above age 40 years from 1980 through 2006 comprising 4 412 568 individuals. Median surveillance time was 23 years. Baseline characteristics are shown in Table 1. We identified 129 206 individuals with

non-melanoma skin cancer, 22 107 with cutaneous malignant melanoma, 327 856 with myocardial infarction, 129 419 with hip fracture and 1 629 519 individuals who died. Mean age of outcomes was 68 years for diagnosis of non-melanoma skin cancer, 59 years for cutaneous malignant melanoma, 69 years for myocardial infarction, 78 years for hip fracture and 76 years for death from any cause.

Myocardial infarction

Cumulative incidence of myocardial infarction as a function of age was lower among individuals with non-melanoma skin cancer (log rank, P -value $< 2 \times 10^{-308}$) and individuals with cutaneous malignant melanoma (log rank, P -value $= 5 \times 10^{-67}$), than among individuals without (Figure 1).

In individuals with vs without non-melanoma skin cancer, the multifactorially adjusted odds ratio was 0.96 (95% confidence interval: 0.94–0.98) for myocardial infarction (Table 2, top). The corresponding odds ratio in individuals with cutaneous malignant melanoma compared with individuals without was 0.79 (0.74–0.84). Stratifying by baseline characteristics only changed odds ratios slightly in most strata (Table 2).

Hip fracture

Cumulative incidence of hip fracture as a function of age was lower among individuals with non-melanoma skin cancer (log rank, P -value $= 9 \times 10^{-233}$) and individuals with cutaneous malignant melanoma (log rank, P -value $= 1 \times 10^{-28}$), than among individuals without (Figure 1).

In individuals with vs without non-melanoma skin cancer, the multifactorially adjusted odds ratio was 1.15 (1.12–1.18) for hip fracture (Table 2, top). The corresponding odds ratio in individuals with cutaneous malignant melanoma compared with individuals without was 0.84 (0.76–0.93). Stratifying by baseline characteristics only changed odds ratios slightly in most strata (Table 2).

The odds ratio of 1.15 (1.12–1.18) for hip fracture in those with vs without non-melanoma skin cancer could be because those with skin cancer live longer and therefore eventually will fall and have a hip fracture, which is particularly common in the elderly. We therefore made a age-stratified analysis and estimated the odds ratio for hip fracture in age-strata of 10 years: up to age 80–89 years the odds ratio was below 1.0, whereas for age-strata 90–99 and >100 years the odds ratios were nominally above 1.0 (Figure 2).

Death from any cause

Cumulative incidence of death from any cause as a function of age was lower among individuals with non-melanoma skin cancer (log rank, P -value $< 2 \times 10^{-308}$) compared with individuals without (Figure 1). Cumulative incidence of death from any

Table 1 Baseline characteristics

	Individuals without skin cancer		Individuals with non-melanoma skin cancer		Individuals with cutaneous malignant melanoma ^b	
	Number	Percent within stratum	Number	Percent within stratum	Number	Percent within stratum
Total	4 261 345		129 206		22 017	
Gender						
Men	2 091 806	49%	64 147	50%	9 696	47%
Women	2 169 539	51%	65 059	50%	12 321	53%
Descent						
Danish	3 988 308	94%	124 819	97%	21 354	97%
Other	261 784	6%	4 372	3%	659	3%
Occupational sun exposure ^a						
Low	3 929 549	96%	118 642	95%	20 362	96%
High	172 451	4%	6 342	5%	928	4%
Residential city size						
<12 000 inhabitants or rural	1 688 871	40%	45 412	35%	8 399	38%
12 100 000 inhabitants	1 092 173	26%	33 574	26%	5 882	27%
>100 000 inhabitants	1 466 266	34%	50 200	39%	7 731	35%
Occupational physical activity ^a						
Low	2 418 807	59%	74 242	59%	11 699	55%
Intermediate	807 832	20%	26 144	21%	5 001	23%
High	890 986	21%	24 709	20%	4 635	22%
Highest level of education						
Unknown ^c	1 042 158	25%	45 722	35%	4 225	19%
Primary school	1 206 611	29%	32 653	25%	6 082	28%
High school	95 348	2%	1 812	1%	418	3%
Vocational education	1 131 858	27%	29 141	23%	6 510	30%
Short academic education	137 917	3%	3 304	2%	770	4%
Medium academic education	405 957	10%	10 927	9%	2 755	13%
Long academic education	175 194	4%	5 549	5%	1 173	5%

Baseline was at study inclusion in 1980, 40th birthday, or at immigration (whichever occurred last). Numbers of individuals vary slightly due to availability of data.

^aFrom 1 January 1980 through 31 December 1995, Statistics Denmark also recorded detailed information on occupation, which allowed us to generate two composite variables: a variable of estimated occupational sun exposure (low or high) and a variable of estimated occupational physical activity (low, intermediate, or high).

^bIndividuals with both non-melanoma skin cancer and cutaneous malignant melanoma were counted only in the cutaneous malignant melanoma group.

^cInformation regarding education was not available if the education was completed prior to 1980 or abroad.

cause as a function of age was higher among individuals with cutaneous malignant melanoma (log rank, P -value = 1×10^{-28}) compared with individuals without, except above age ~ 70 years (Figure 1).

In individuals with vs without non-melanoma skin cancer, the multivariately adjusted hazard ratio was 0.52 (0.52–0.53) for death from any cause (Table 2, top). The corresponding hazard ratio in individuals with cutaneous malignant melanoma compared with individuals without was 0.89 (0.87–0.91). Stratifying

by baseline characteristics only changed hazard ratios slightly in most strata (Table 2). For cutaneous malignant melanoma, stratifying by level of education showed reduced risk of death from any cause in the group with unknown educational level (dominated by older people), and higher risk of death from any cause in those with high school or more advanced educational level, largely reflecting domination by younger individuals in these latter groups. This pattern was also seen in the age-stratified analysis: among individuals

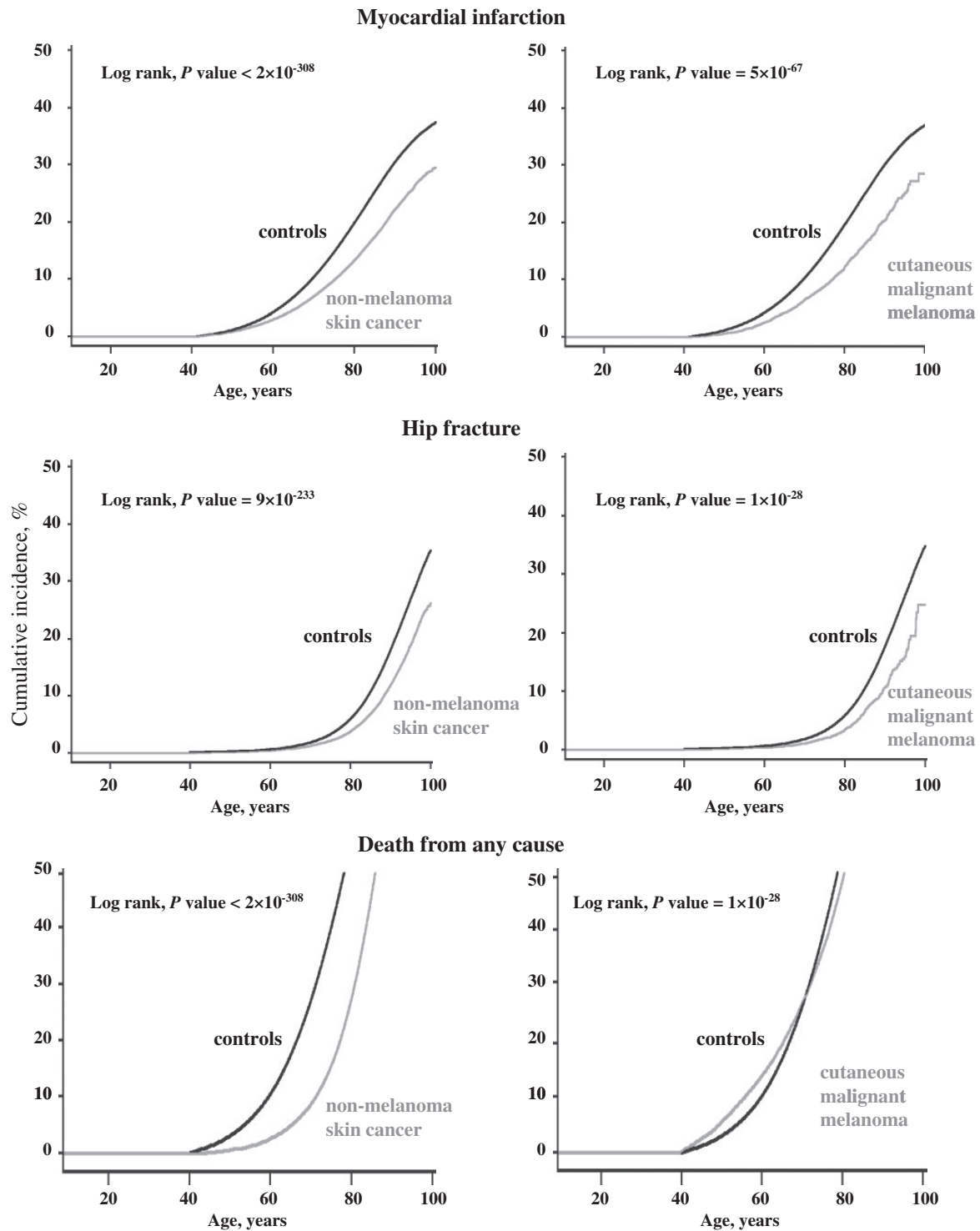


Figure 1 The cumulative incidence of myocardial infarction, hip fracture and death as a function of age in individuals above age 40 years ever diagnosed with non-melanoma skin cancer and cutaneous malignant melanoma. Cumulative incidence curves were generated from Kaplan–Meyer estimates, comparing individuals with non-melanoma skin cancer and cutaneous malignant melanoma vs individuals free of both diseases. P -values are for comparison between groups by log rank tests

Table 2 Odds ratios of myocardial infarction and hip fracture, and hazard ratios of death from any cause, in individuals ever diagnosed with non-melanoma skin cancer or cutaneous malignant melanoma in the entire Danish population above age 40 years stratified by baseline characteristic

	Non-melanoma skin cancer			Cutaneous malignant melanoma		
	Myocardial infarction OR (95% CI)	Hip fracture OR (95% CI)	Death from any cause HR (95% CI)	Myocardial infarction OR (95% CI)	Hip fracture OR (95% CI)	Death from any cause HR (95% CI)
All	0.96 (0.94–0.98)	1.15 (1.12–1.18)	0.52 (0.52–0.53)	0.79 (0.74–0.84)	0.84 (0.76–0.93)	0.89 (0.87–0.91)
Gender						
Men	0.96 (0.94–0.98)	1.17 (1.11–1.22)	0.52 (0.51–0.53)	0.76 (0.70–0.82)	0.79 (0.65–0.95)	0.90 (0.87–0.93)
Women	0.96 (0.93–0.99)	1.15 (1.11–1.19)	0.53 (0.53–0.54)	0.82 (0.75–0.91)	0.87 (0.78–0.98)	0.89 (0.86–0.92)
Descent						
Danish	0.96 (0.94–0.98)	1.14 (1.11–1.18)	0.52 (0.52–0.53)	0.79 (0.74–0.84)	0.85 (0.77–0.93)	0.89 (0.87–0.91)
Other	1.04 (0.93–1.17)	1.36 (1.15–1.60)	0.55 (0.52–0.58)	0.86 (0.60–1.23)	0.59 (0.27–1.26)	1.03 (0.90–1.17)
Occupational sun exposure						
Low	0.97 (0.95–0.99)	1.17 (0.99–1.38)	0.52 (0.52–0.53)	0.80 (0.75–0.85)	0.84 (0.76–0.93)	0.88 (0.86–0.90)
High	0.89 (0.83–1.06)	1.15 (1.11–1.18)	0.58 (0.55–0.60)	0.68 (0.54–0.86)	0.80 (0.45–1.42)	1.03 (0.94–1.13)
Residential city size						
<12 000 inhabitants or rural	0.94 (0.91–0.97)	1.15 (1.09–1.20)	0.55 (0.54–0.56)	0.71 (0.64–0.85)	0.91 (0.78–1.07)	0.97 (0.93–1.00)
12–100 000 inhabitants	0.96 (0.92–1.00)	1.14 (1.08–1.21)	0.53 (0.52–0.54)	0.84 (0.75–0.94)	0.78 (0.64–0.96)	0.90 (0.86–0.94)
>100 000 inhabitants	0.98 (0.95–1.01)	1.15 (1.10–1.20)	0.50 (0.50–0.51)	0.84 (0.76–0.93)	0.82 (0.70–0.96)	0.82 (0.79–0.84)
Occupational physical activity						
Low	0.99 (0.83–0.91)	1.15 (1.11–1.18)	0.55 (0.54–0.57)	0.84 (0.77–0.91)	0.84 (0.75–0.94)	1.01 (0.97–1.06)
Intermediate	0.85 (0.81–0.89)	1.03 (0.94–1.13)	0.52 (0.51–0.54)	0.66 (0.58–0.76)	0.68 (0.50–0.93)	1.04 (0.99–1.10)
High	0.87 (0.83–0.96)	1.06 (0.97–1.15)	0.52 (0.51–0.52)	0.73 (0.65–0.82)	0.88 (0.69–1.12)	0.81 (0.79–0.84)
Highest level of education						
Unknown ^a	0.91 (0.90–0.94)	1.14 (1.10–1.18)	0.52 (0.52–0.53)	0.73 (0.66–0.81)	0.86 (0.76–0.97)	0.71 (0.68–0.73)
Primary school	0.82 (0.79–0.86)	0.90 (0.83–0.97)	0.52 (0.51–0.53)	0.72 (0.64–0.80)	0.74 (0.59–0.91)	0.98 (0.94–1.02)
High school	0.86 (0.68–1.08)	0.91 (0.61–1.38)	0.49 (0.43–0.55)	1.05 (0.62–1.77)	0.90 (0.32–2.46)	1.38 (1.13–1.67)
Vocational education	0.81 (0.77–0.85)	0.95 (0.85–1.05)	0.51 (0.50–0.52)	0.71 (0.62–0.81)	0.64 (0.46–0.88)	1.15 (1.10–1.21)
Short academic education	0.93 (0.80–1.09)	0.75 (0.52–1.08)	0.53 (0.49–0.58)	0.58 (0.36–0.93)	0.90 (0.37–2.19)	1.46 (1.25–1.71)
Medium academic education	0.76 (0.68–0.84)	1.00 (0.82–1.22)	0.53 (0.50–0.56)	0.60 (0.46–0.80)	0.51 (0.27–0.99)	1.40 (1.28–1.52)
Long academic education	0.94 (0.83–1.06)	1.08 (0.84–1.39)	0.57 (0.54–0.61)	0.74 (0.53–1.05)	0.69 (0.31–1.57)	1.24 (1.10–1.40)

OR, odds ratio; HR, hazard ratio; CI, confidence interval.

Odd ratios are from logistic regression analysis and hazard ratios are from Cox regression analysis, both including the entire population above age 40 years and adjusted multivariate for age, gender, descent, occupational sun exposure, residential city size, occupational physical activity and highest level of education.

^aInformation regarding education was not available if the education was completed prior to 1980 or abroad.

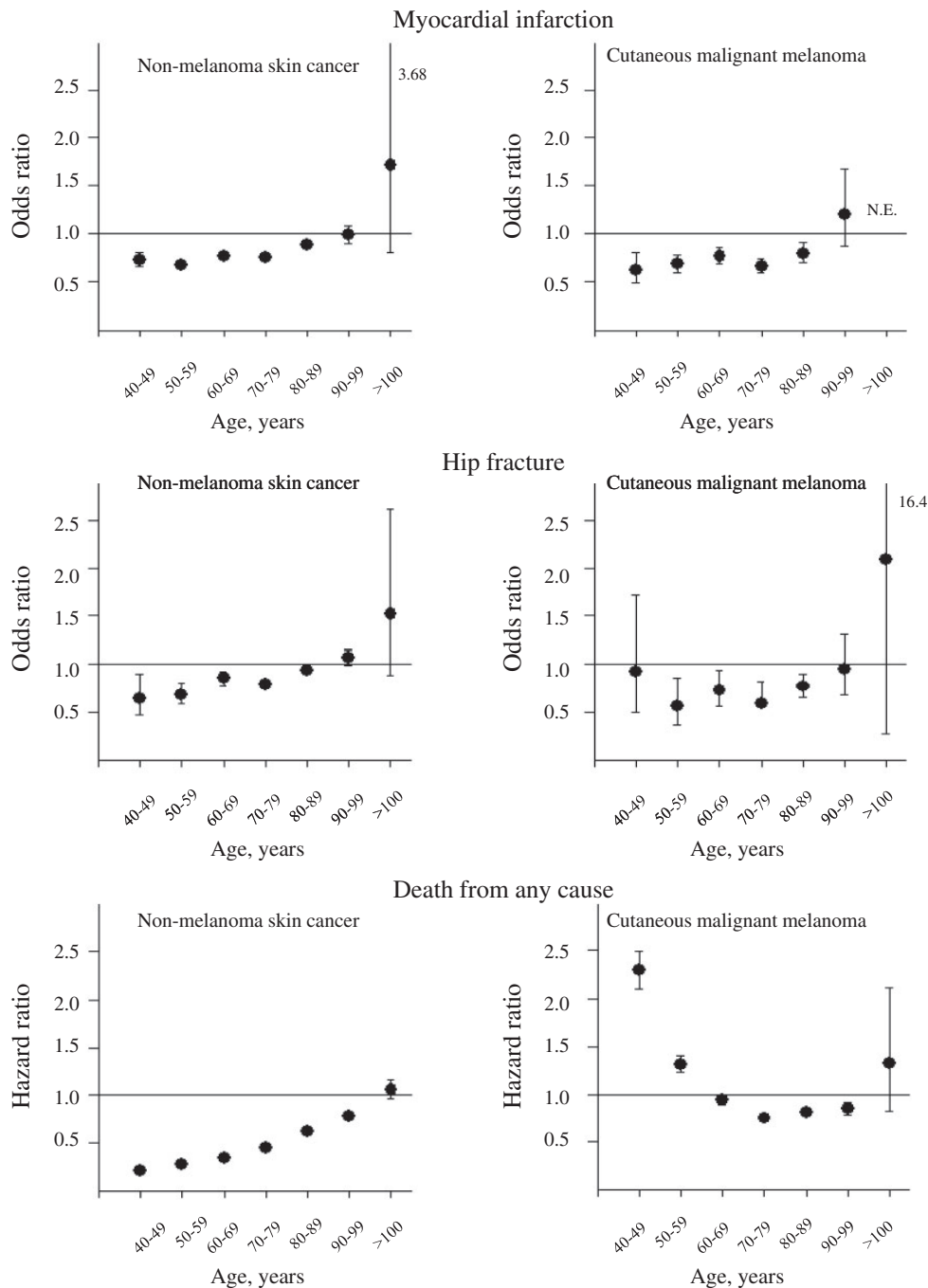


Figure 2 In the entire Danish population above age 40 years, odds ratios for myocardial infarction and hip fracture and hazard ratios for death from any cause within 10-years age-strata. N.E., no estimation due to limited statistical power

aged 40–59 years there was an increased risk of death from any cause, whereas this was not the case for individuals at age 60 and above (Figure 2).

Birth year-, age- and gender-matched case-control study

To circumvent the effect of time (calendar year), changes in sun exposure habits and changes in treatment of cancer during the observation period, we also

examined the risk of myocardial infarction, hip fracture and death from any cause in individuals with non-melanoma skin cancer or cutaneous malignant melanoma matched with five general population controls on birth year, age and gender. For these analyses only myocardial infarction and hip fracture events following a diagnosis of non-melanoma skin cancer or cutaneous malignant melanoma entered into the analysis, whereas events before skin cancer were excluded.

Table 3 Odds ratios of myocardial infarction, hip fracture and death from any cause in individuals above 40 years with non-melanoma skin cancer and cutaneous malignant melanoma

	Non-melanoma skin cancer		Cutaneous malignant melanoma	
	Age-adjusted OR (95% CI)	Multifactorially adjusted OR (95% CI)	Age-adjusted OR (95% CI)	Multifactorially adjusted OR (95% CI)
Myocardial infarction	0.87 (0.85–0.90)	0.90 (0.88–0.92)	0.72 (0.66–0.78)	0.74 (0.68–0.81)
Hip fracture	0.97 (0.92–1.00)	0.99 (0.95–1.02)	0.70 (0.61–0.79)	0.71 (0.62–0.81)
Death from any cause	0.96 (0.95–0.97)	0.97 (0.96–0.99)	1.55 (1.50–1.60)	1.96 (1.89–2.04)

OR, odds ratio; CI, confidence interval.

To circumvent the effect of time (calendar year), changes in sun exposure habits and change in treatment of cancer during the past three decades, individuals with non-melanoma skin cancer or cutaneous malignant melanoma were each matched with five general population controls of the same birth year, age and gender. Birth year was matched beside age to also take into account that people born at different time periods throughout history have used sun exposure to a different degree.

In individuals with vs without non-melanoma skin cancer, the multifactorially adjusted odds ratios were 0.90 (0.88–0.99) for myocardial infarction, 0.99 (0.95–1.02) for hip fracture and 0.97 (0.96–0.99) for death from any cause (Table 3). In individuals with vs without cutaneous malignant melanoma, the multifactorially adjusted odds ratios were 0.74 (0.68–0.81) for myocardial infarction, 0.71 (0.62–0.81) for hip fracture, and 1.96 (1.89–2.04) for death from any cause (Table 3). In sensitivity analyses, corresponding odds ratios in 10-year age strata are shown in Figure 3.

Discussion

In a nationwide study of 4.4 million individuals above age 40 years, having a diagnosis of skin cancer was associated with less myocardial infarction, less hip fracture in those below age 90 years and less death from any cause. However, cutaneous malignant melanoma was associated positively with death from any cause in some individuals. As skin cancer is a marker of a substantial sun exposure, these results indirectly suggest that sun exposure might have beneficial effects on health. However, causal or mechanistic conclusions cannot be drawn from this study design and a potential beneficial effect of sun exposure *per se* needs to be examined in other studies.

Mechanistically, one could however speculate that our findings theoretically could be explained by an association between increased sun exposure and more outdoor physical activity. In accordance with this, there is an inverse linear dose-response between physical activity and risk of cardiovascular disease, osteoporosis and all-cause mortality.^{8,9} In further support of this idea are the findings in the present study of the lowest risk of myocardial infarction and death from any cause in individuals with a high level of occupational physical activity.

Another theoretically possible explanation of our findings relates to the fact that increased sun exposure also associates with increased vitamin D synthesis. Vitamin D exerts both direct and indirect

endocrine, immunomodulatory and neurohormonal effects on the cells of the cardiovascular system, potentially leading to an overall protection against cardiovascular disease.^{10–12} An association between high levels of vitamin D and lower cardiovascular morbidity and mortality has been reported in several epidemiological studies,¹³ whereas randomized controlled trials show no effect of supplementation with vitamin D on risk of cardiovascular mortality.¹⁴ Reports on vitamin D and risk of osteoporosis are ambiguous; results from epidemiological studies show that high levels of vitamin D associate with decreased risk of hip fracture, whereas results from meta-analyses of randomized controlled trials have failed to show an effect on risk of hip fracture.^{15,16} However, a meta-analysis of vitamin D and calcium supplementation combined concludes that this treatment lowers risk of hip fracture.¹⁷ The association between high levels of vitamin D and lower mortality has been demonstrated both in epidemiological studies and in several randomized controlled trials with mortality as a secondary outcome.^{18,19}

In the present study, age is a potential effect modifier; to address this possibility we have restricted all analyses to age above 40 years, adjusted for age in the logistic regression analyses, in the Cox regression analysis used age as the underlying intensity and in the matched study matched on age. Moreover, we have performed age-stratified analyses in age-strata of 10 years. For hip fracture, although the overall odds ratio was 1.15 (95% CI: 1.12–1.18), among those below age 90 years odds ratios in those with vs without non-melanoma skin cancer were below 1.0. This suggests that individuals with non-melanoma skin cancer, which is most often a benign condition, sometimes live longer with a consequent increase in risk of hip fracture at very old age. Although, in our data, individuals with cutaneous malignant melanoma and high occupational sun exposure showed no association with mortality from any cause, including from cancer, it has been shown that sun exposure may increase survival from malignant melanoma; this suggests that cutaneous malignant melanoma may be

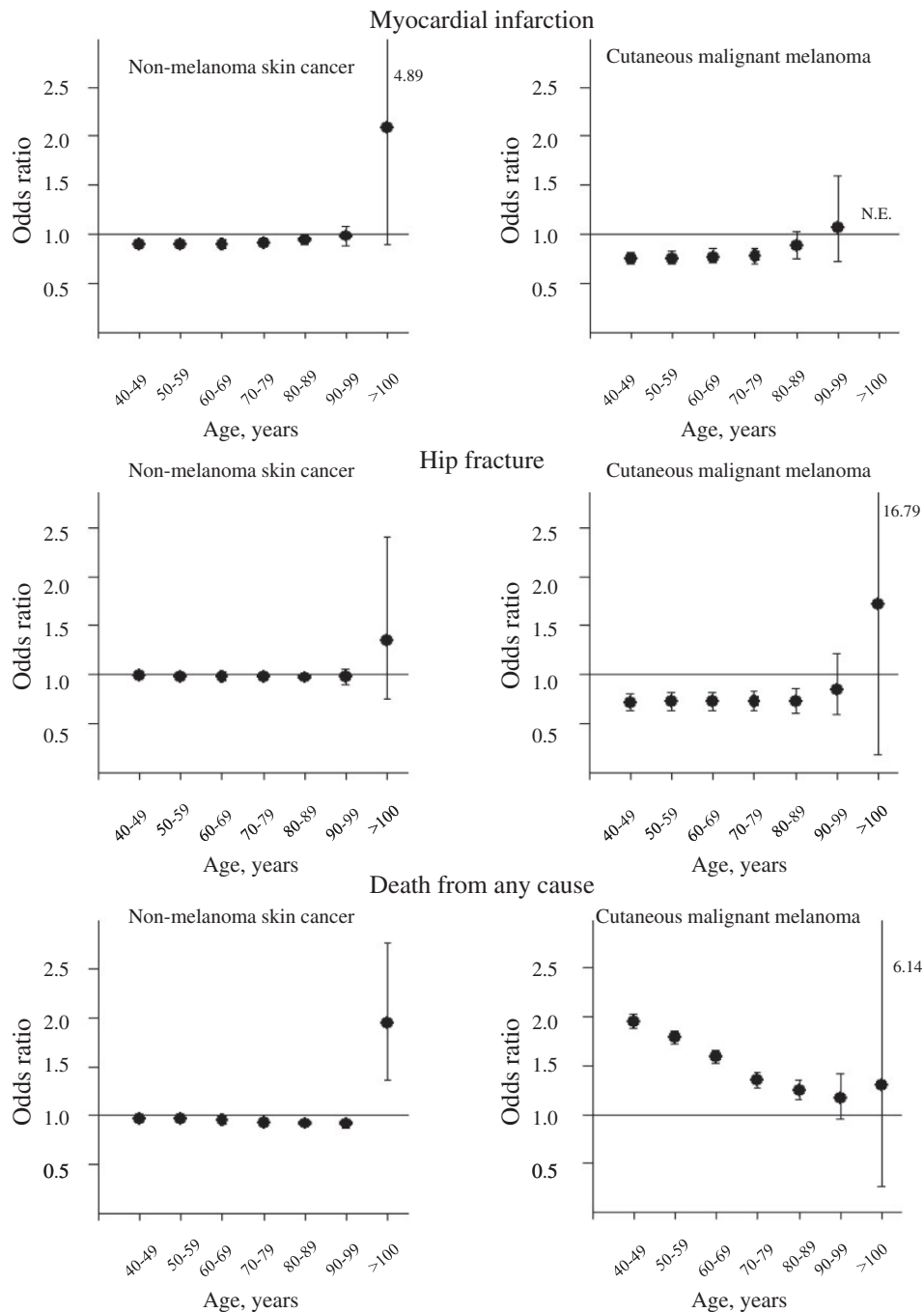


Figure 3 In a matched study within the entire Danish population above age 40 years, odds ratios for myocardial infarction, hip fracture and death from any cause within 10-years age-strata. Individuals with non-melanoma skin cancer or cutaneous malignant melanoma were each matched with five general population controls of the same birth year, age and gender. N.E., no estimation due to limited statistical power

biologically more benign if it occurs in association with high levels of sun exposure.²⁰ The difference in estimates between individuals with a diagnosis of non-melanoma skin cancer and individuals with a diagnosis of cutaneous malignant melanoma could be due to the fact that non-melanoma skin cancer is most often a benign condition, and thus individuals

with this diagnosis live longer and have 'the full benefit' of sun exposure throughout life, as opposed to individuals with cutaneous malignant melanoma, who often die early.

A strength of the present study is the use of a large nationwide cohort, with complete registration of diagnoses, death and migration and with a median

follow-up time of 23 years. Limitations include that the study population mostly consists of Whites and results may therefore not necessarily apply to other ethnic groups. Also, a limitation of the use of skin cancer diagnoses as a proxy for sun exposure is that not all skin cancers are caused by sun exposure. This presumably smaller fraction of skin cancers could have their own association to the outcomes studied, and may therefore cause both under- and overestimation of the observed associations. Moreover, there is a large variability in the genetic susceptibility to development of skin cancer upon ultraviolet radiation exposure,²¹ and therefore some individuals with very high level of sun exposure may not develop skin cancer whereas other individuals with low levels of sun exposure develop skin cancer. This would in both cases lead to an attenuation of the observed associations, and therefore cannot explain the present findings. Furthermore, for the outcome myocardial infarction, our results could be biased by cases of silent myocardial infarction leading to differential misclassification: it is not unlikely that cases of silent myocardial infarction would be more frequently registered in cancer patients in contact with the Danish health care service, and this could either overestimate or underestimate the true association. Finally, a limitation is that we did not have information from Statistics Denmark on smoking status and we cannot adjust for smoking status in our analyses; smoking could be an important confounder or effect modifier associated with exposures as well as outcome variables.

In conclusion, the present study suggests that having a diagnosis of skin cancer was associated with less myocardial infarction, less hip fracture in those below age 90 years and less death from any cause compared with general population controls.

Although some individuals with cutaneous malignant melanoma experience increased risk of death from any cause, the overall data indirectly suggest that sun exposure for many individuals may have beneficial health effects, and therefore also question the widespread advice that sun exposure should be avoided. Nevertheless, a potential beneficial effect of sun exposure per se needs to be examined in other studies.

Funding

This work was supported by the Danish Heart Foundation [10-01-R79-A2793-22574], the Faculty of Health Sciences, University of Copenhagen, and by Herlev Hospital, Copenhagen University Hospital. The study funders had no role in the design or conduct of the study; in the collection, analysis, management and interpretation of the data; or in the preparation, review or approval of the manuscript. All authors are independent from funders.

Acknowledgements

B.G.N. initiated the study, which was designed in detail by P.B.J., B.G.N., S.F.N. and M.B. All authors had full access to all of the data. Database handling and statistical analyses were by P.B.J., B.G.N., S.F.N. and M.B.. All four authors contributed to analyses and interpretation of data. P.B.J. wrote the first draft of the paper, which was revised and finally accepted by the other three authors.

Conflict of interest: None declared.

KEY MESSAGES

- In a nationwide study of 4.4 million individuals above age 40 years, having a diagnosis of skin cancer was associated with less myocardial infarction, less hip fracture and less death from any cause.
- As skin cancer is a marker of a substantial sun exposure, these results indirectly suggest that sun exposure might have beneficial effects on health.
- However, causal or mechanistic conclusions cannot be drawn from this study design and a potential beneficial effect of sun exposure per se needs to be examined in other studies.

References

- ¹ Reichrath J. The challenge resulting from positive and negative effects of sunlight: how much solar UV exposure is appropriate to balance between risks of vitamin D deficiency and skin cancer? *Prog Biophys Mol Biol* 2006;**92**: 9–16.
- ² Jensen AO, Bautz A, Olesen AB, Karagas MR, Sorensen HT, Friis S. Mortality in Danish patients with non-melanoma skin cancer, 1978-2001. *Br J Dermatol* 2008;**159**:419–25.
- ³ Armstrong BK, Kricger A. The epidemiology of UV induced skin cancer. *J Photochem Photobiol B* 2001;**63**: 8–18.
- ⁴ Kenborg L, Jorgensen AD, Budtz-Jorgensen E, Knudsen LE, Hansen J. Occupational exposure to the sun and risk of skin and lip cancer among male wage earners in Denmark: a population-based case-control study. *Cancer Causes Control* 2010;**21**:1347–55.
- ⁵ Madan V, Lear JT, Szeimies RM. Non-melanoma skin cancer. *Lancet* 2010;**375**:673–85.

- ⁶ Berwick M, Armstrong BK, Ben-Porat L *et al*. Sun exposure and mortality from melanoma. *J Natl Cancer Inst* 2005; **97**:195–99.
- ⁷ Afzal S, Nordestgaard BG, Bojesen SE. Plasma 25-hydroxyvitamin D and risk of non-melanoma and melanoma skin cancer: a prospective cohort study. *J Invest Dermatol* 2013; **133**:629–36.
- ⁸ Haennel RG, Lemire F. Physical activity to prevent cardiovascular disease. How much is enough? *Can Fam Physician* 2002; **48**:65–71.
- ⁹ Schmitt NM, Schmitt J, Doren M. The role of physical activity in the prevention of osteoporosis in postmenopausal women – An update. *Maturitas* 2009; **63**:34–38.
- ¹⁰ Nemerovski CW, Dorsch MP, Simpson RU, Bone HG, Aaronson KD, Bleske BE. Vitamin D and cardiovascular disease. *Pharmacotherapy* 2009; **29**:691–708.
- ¹¹ Takeda M, Yamashita T, Sasaki N *et al*. Oral administration of an active form of vitamin D3 (calcitriol) decreases atherosclerosis in mice by inducing regulatory T cells and immature dendritic cells with tolerogenic functions. *Arterioscler Thromb Vasc Biol* 2010; **30**:2495–503.
- ¹² Li YC, Kong J, Wei M, Chen ZF, Liu SQ, Cao LP. 1,25-Dihydroxyvitamin D(3) is a negative endocrine regulator of the renin-angiotensin system. *J Clin Invest* 2002; **110**:229–38.
- ¹³ Parker J, Hashmi O, Dutton D *et al*. Levels of vitamin D and cardiometabolic disorders: Systematic review and meta-analysis. *Maturitas*, 2010; **65**:225–36.
- ¹⁴ Judd SE, Tangpricha V. Vitamin D deficiency and risk for cardiovascular disease. *Am J Med Sci* 2009; **338**:40–44.
- ¹⁵ Lai JK, Lucas RM, Clements MS, Roddam AW, Banks E. Hip fracture risk in relation to vitamin D supplementation and serum 25-hydroxyvitamin D levels: a systematic review and meta-analysis of randomised controlled trials and observational studies. *BMC Public Health* 2010; **10**:331.
- ¹⁶ Avenell A, Gillespie WJ, Gillespie LD, O’Connell D. Vitamin D and vitamin D analogues for preventing fractures associated with involutional and post-menopausal osteoporosis. *Cochrane Database Syst Rev* 2009; CD000227.
- ¹⁷ Patient level pooled analysis of 68 500 patients from seven major vitamin D fracture trials in US and Europe. *BMJ* 2010; **340**:b5463.
- ¹⁸ Zittermann A, Gummert JF, Bergermann J. Vitamin D deficiency and mortality. *Curr Opin Clin Nutr Metab Care* 2009; **12**:634–39.
- ¹⁹ Autier P, Gandini S. Vitamin D supplementation and total mortality: a meta-analysis of randomized controlled trials. *Arch Intern Med* 2007; **167**:1730–37.
- ²⁰ Berwick M, Armstrong BK, Ben-Porat L *et al*. Sun exposure and mortality from melanoma. *J Natl Cancer Inst* 2005; **97**:195–99.
- ²¹ Madan V, Lear JT, Szeimies RM. Non-melanoma skin cancer. *Lancet* 2010; **375**:673–85.



Don't laugh, he's going to live longer than you do (Photo: Coloubox)

October 16, 2013
08:00

by KM

Sunbathers live longer

Spending time in the sun can add years to your life, a 20-year study following the health of 4.4 million Danes finds. The team of Danish scientists, whose research results will be published in the Journal of Epidemiology, found that people who were regular sunbathers and who had developed benign forms of skin cancer lived up to six years longer than the average for the population as a whole. The study also found that sunbathers had lower rates of heart attacks and osteoporosis. While the team said its evidence was conclusive, they said they had not been unable to determine what made sunbathers live longer. – *Politiken*

SEE RELATED: More Danes dying of cancer

PM, opposition leader now in dead heat

For the first time since the 2011 general election, Lars Løkke Rasmussen (Venstre), the opposition leader, has lost his lead over the prime minister in the polls. After two weeks of bad press, first after over-estimating the cost of a price of shoes, then for travelling first-class at tax-payer expense, Rasmussen's support has shrunk to 37 percent, a loss of 10 percentage points. Meanwhile PM Helle Thorning-Schmidt has made up significant ground, seeing her approval ratings rise seven percentage points to 39 percent. Rasmussen's lieutenants expected he would bounce back, but political analysts warned Venstre against expecting the issue would disappear on its own. "This is dangerous, because we're not talking about a single slip-up. It is reminiscent of previous problems he had with being repaid for unjustified expenses," said Rune Stubager, Aarhus University. – *Berlingske*

SEE RELATED: Right wing surge confirmed

State could open its gates to foreign entrepreneurs

Non-western immigrants live longer

Diner fined for whining, but did café cross the line?

Latest Comments

Also I believe a lot of these immigrants are trying to imitate the Danes by...
(Hamish Carey on April 7, 2014 17:51)

He expects 50 permits be issued under the scheme during the first year, with...
(Jon Paris King on April 7, 2014 16:42)

What do you expect from a system that supports wrong in all it's form because...
(Alæen-Roger Mbah on April 7, 2014 14:28)

?You need to go to a ghetto in the USA or a Brazilian favela to see something...
(Leo Carona on April 7, 2014 08:52)

Since there has never been a world wide flood who cares how they portray a myth.
(Lewis Thomason on April 7, 2014 00:20)

About the fork and knife, I think means an "open" sandwich?
(Jens Rost on April 6, 2014 17:19)

One of my favorite Danish immigrants to the USA was George "Dutch" Anderson...
(Bill Jones on April 6, 2014 16:20)

Yeah, this is truly a valid reason for hoisting the flag. For once, I'd not...
(Abhijeet Mishra on April 6, 2014 15:41)

Now dont show off with your big car boy . !
(Hamish Carey on April 6, 2014 10:05)

The theglobaledition.com couldn't wait until April 1 to run a satirical story...
(Corgi Hill Farm on April 6, 2014 02:52)

Focus On



Istedgade makeover casts red light district in a new light

New rules drive up the cost of love