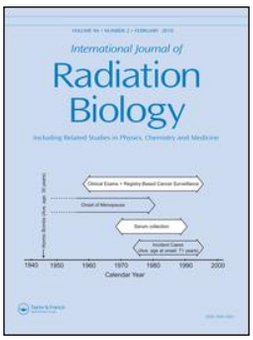

International Journal of Radiation Biology

ISSN: 0955-3002 (Print) 1362-3095 (Online) Journal homepage: http://www.tandfonline.com/loi/irab20

Evaluation of statistical modeling approaches for epidemiologic studies of low-dose radiation health effects

Ashley P. Golden, Sarah S. Cohen, Heidi Chen, Elizabeth D. Ellis & John D. Boice Jr.

To cite this article: Ashley P. Golden, Sarah S. Cohen, Heidi Chen, Elizabeth D. Ellis & John D. Boice Jr. (2018): Evaluation of statistical modeling approaches for epidemiologic studies of low-dose radiation health effects, International Journal of Radiation Biology, DOI: 10.1080/09553002.2018.1554924

To link to this article: https://doi.org/10.1080/09553002.2018.1554924

INTERNATIONAL JOURNAL OF RADIATION BIOLOGY
https://doi.org/10.1080/09553002.2018.1554924

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Evaluation of statistical modeling approaches for epidemiologic studies of low-dose radiation health effects

Ashley P. Golden[a], Sarah S. Cohen[b], Heidi Chen[c], Elizabeth D. Ellis[a] and John D. Boice, Jr.[c,d]

[a]Oak Ridge Associated Universities, Oak Ridge, TN, USA; [b]EpidStat Institute, Ann Arbor, MI, USA; [c]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center and Vanderbilt-Ingram Cancer Center, Nashville, TN, USA; [d]National Council on Radiation Protection and Measurements, Bethesda, MD, USA

ABSTRACT

Purpose: A substantial body of epidemiologic literature addresses risks associated with occupational radiation exposure but comparing results between studies is often difficult as different statistical models are commonly used. It is unclear whether different methods produce similar results for estimates of radiation risk when applied to the same data. The goal of this study was to compare the radiation risk estimates for leukemia other than chronic lymphocytic leukemia (non-CLL) and ischemic heart disease (IHD) produced by both Cox and Poisson regression models for time-dependent dose-response analyses of occupational exposure.

Materials and methods: For brevity, this methods paper presents the results from one cohort, the Nuclear Power Plant workers (NPP), though the evaluation considered five cohorts of varying size and exposure as part of the Million Worker Study. Cox Proportional Hazards models, with age as the underlying timescale for hazard, were conducted using three computer software programs: SAS, R, and Epicure. Doses lagged 2 years for non-CLL and 10 years for ischemic heart disease were treated as time-dependent exposures at the annual level and were examined both in categories and as a continuous term. Hazard ratios (HR) and 95% confidence intervals (CI) were reported for each model in SAS and R, while the Peanuts program of Epicure was utilized to produce Excess Relative Risk (ERR) estimates and 95% CI. All models were adjusted for gender and year of birth. Four piece-wise exponential Poisson models (log-linear regression for rate) were developed with varying cutpoints for age strata from very fine to broad categories using both R and the Amfit program in Epicure for ERR estimates.

Results: Comparable estimates of risk (both RR and ERR) were observed from Cox and Poisson models, regardless of software utilized, as long as appropriately narrow categories of age were utilized to control the confounding of age in Poisson models. The ERR estimates produced in Epicure tended to agree very closely with the HR or RR estimates, and the statistical software program used had no impact to risk estimates for the same model.

Conclusions: As computational power is no longer the burden today as it has been in the past, the results of this evaluation support the use of the Cox proportional hazards or the ungrouped Poisson approach to analyzing time-dependent dose-response relationships to ensure that maximum control over the confounding of age is achieved in studies of mortality for cohorts occupationally exposed to radiation.

ARTICLE HISTORY

Received 17 September 2018
Revised 27 November 2018
Accepted 27 November 2018

KEYWORDS

Epidemiology; models of radiation responses; statistical models; Cox proportional hazards model; Poisson regression model

Introduction

While the studies of the atomic bomb survivors (Preston et al. 2007; Yamada et al. 2009; Shimizu et al. 2010; Ozasa et al. 2012; Hsu et al. 2013; Sera et al. 2013; Ozasa et al. 2017) have provided estimates for health risks resulting from exposure to ionizing radiation, the acute exposure received by this population is not characteristic of the chronic exposure to low-dose ionizing radiation encountered more broadly in the occupational setting. The body of epidemiologic literature addressing this workplace risk has grown significantly in the last 10–15 years (Pinkerton et al. 2004; Boice et al. 2006; McGeoghegan and Binks 2006; Boice et al. 2007, 2008, 2011, 2014; Cardis et al. 2007; Hornung et al. 2008; Schubauer-Berigan et al. 2009; Yiin et al. 2009, 2017, 2018 2018; Guseva Canu et al. 2010; Anderson et al. 2012; Eidemuller et al. 2012; Guseva Canu et al. 2012; Silver et al. 2013; Zablotska et al. 2013, 2014, 2018; Kreuzer et al. 2015; Zhivin et al. 2016; Richardson et al. 2018; Shore et al. 2018). In reviewing these studies, there are a number of important methodologic issues that make comparison of results difficult. Two different statistical approaches are commonly used in occupational studies of radiation exposure: Cox proportional hazards modeling (Cox 1972) and Poisson piece-wise exponential regression (Frome 1983, 1985). Cox models produce hazards ratios (HR) while Poisson regression results are presented as relative risk (RR). Both methods can be utilized to estimate excess relative risk (ERR) or excess absolute risk (EAR).

CONTACT Ashley Golden ashley.golden@orau.org 100 ORAU Way, Oak Ridge, TN 37831, USA.

Supplemental data for this article can be accessed here.

Poisson regression has been viewed as an extension of an age-standardized rate ratio (such as a Standardized Mortality Ratio, SMR, for mortality studies), where the data are typically cross-classified by levels of exposure and other covariates, and person-years grouped for each resulting combination or cell (Breslow and Day 1987). Thus, changes in exposure or other covariates, including age, over time are taken into account by changing the covariate pattern to which the person-years are assigned, which is referred to as 'stratified', 'grouped', or 'piece-wise exponential' Poisson regression. The Poisson model assumes that the rate ratio, or relative risk, among the exposed and non-exposed groups is constant within exposure-age stratum (Checkoway et al. 2004; Allison 2010). As death rates typically rise with age (Breslow et al. 1983) this assumption may not be met if the intervals of time are wide (Callas et al. 1998; Richardson and Loomis 2004; Allison 2010).

As age is often a strong confounder for mortality studies, it is desirable to obtain maximum possible control over the time-dependent effects by stratifying very finely on age (Checkoway et al. 2004). This is achieved such that every age-stratum would contain at most one death through the specification of Cox proportional hazards model (Cox 1972) as an extension of Poisson regression, rather than as a special case of survival analysis (Pearce et al. 1988). Unlike Poisson regression, the age-strata are not included as covariates in the model; rather, age is the underlying timescale that defines the persons at risk at a given time (age) for each case of death (Allison 2010). Thus, the calculated rate ratio is assumed to be independent of time (age) (Checkoway et al. 2004), such that the Cox model does not require any assumption to be made about the reference population's death rate, or baseline hazard (Breslow 1974; Cox and Oakes 1984; Allison 2010).

While it is known that the Poisson model will converge to the Cox proportional hazards model as the age strata are made infinitely small (Checkoway et al. 2004), and several studies of empirical comparisons have demonstrated comparable results between the two models (Ingram and Kleinman 1989; Callas et al. 1994, 1998; Loomis et al. 2005), it is nevertheless suggested that researchers should assess the sensitivity of study findings on the analytic method chosen by comparing findings from different approaches (Breslow and Day 1987; Ingram and Kleinman 1989). Additionally, Poisson modeling has most commonly been done using Epicure software (Preston et al. 2015) in the radiation epidemiology literature, but multiple software packages have been used for Cox modeling, including: The SAS System, R, and Epicure. While it would be expected that the software packages produce the same results for a given model in a study population, there does not appear to be any published evaluation of whether the capabilities for time-dependent modeling of exposure and calculations of ERR are comparable, though it is recognized that some authors may perform this as a sensitivity analyses without necessarily reporting the results.

Thus, the goal of this study was to compare the radiation risk estimates for leukemia other than chronic lymphocytic leukemia (non-CLL) and ischemic heart disease (IHD) produced by both Cox and Poisson regression models in several occupationally-exposed cohorts of varying size and exposure as part of the study of a *Million U.S. Radiation Workers and Atomic Veterans* (MWS) (Boice et al. 2018). For brevity, this methods paper presents the results of the evaluation of one cohort, the Nuclear Power Plant workers (NPP). As the intention of this evaluation was to focus on the possible differences in regression methods and associated software, simple models were defined with only a basic set of covariates which may not represent the biologic relevance of all covariates. Thus, the results presented here are not necessarily recommended for use in precisely defining the radiation risk and dose response relationship, but rather are for analytical comparison.

Methods

Study cohorts

Five occupational cohorts from within the MWS were evaluated and are described elsewhere: Atomic Veterans (AV) (Boice, Till, et al. In press); Industrial Radiographers (IR) (Boice et al. In press); Mallinckrodt Chemical Works (MCW) (Dupree-Ellis et al. 2000; Ellis et al. 2018; Golden et al. In press); Mound workers (Boice et al. 2014); and Nuclear Power Plant workers (NPP) (Boice et al. In press);

As a first consideration, the type of radiation exposure varied among the cohorts. The NPP and IR cohorts were primarily exposed to external ionizing radiation, while the Mound and MCW workers had potential for exposure to internal as well as external radiation. Unlike the other cohorts, the Atomic Veterans were present at above-ground nuclear tests resulting in an acute relatively low radiation exposure. Second, the cohorts varied in size. The demographic characteristics of the NPP cohort are presented in Table 1, while characteristics of the other cohorts are provided in Supplemental Table 1. With 145,209 workers, NPP is the largest cohort followed by IR (123, 556), AV (113, 807), Mound (4954), and MCW (2514). The IR cohort has the most females (12,946) compared to NPP (4,975), and Mound (971). MCW and AV are all male cohorts. NPP was chosen as the representative cohort for this paper as it is large, with a substantial number of females, and slightly older than IR but younger than the others when comparing year of birth categories and average length of follow-up.

Vital status ascertainment, cause of death determination, and radiation dose reconstruction were done similarly for all five cohorts. The reconstruction of radiation organ doses for each of these cohorts was done following the guidance outlined in the compendium from the National Commission on Radiation Protection and Measurements entitled 'Deriving Organ Doses and Their Uncertainty for Epidemiologic Studies—Guidance for the One Million U.S. Persons Study of Low-Dose Radiation health Effects' (Bouville et al. 2015; Dauer et al. 2018). Deaths were obtained via extensive tracing as described in detail elsewhere (Mumma et al. 2018). Briefly, linkages to the Social Security Administration death master file as well as state mortality records using the Centers for Disease Control and Prevention LinkPlus program,

Table 1. Characteristics of the nuclear power plant workers cohort.

	Nuclear power plant workers	
	N	%
Sex		
Male	140,234	97%
Female	4975	3%
Year of birth categories		
<1916	1156	1%
1916–1920	2514	2%
1921–1925	4951	3%
1926–1930	6941	5%
1931–1935	9109	6%
1936–1940	12,806	9%
1941–1945	19,391	13%
1946–1950	28,166	19%
1951–1955	30,270	21%
>1955	29,905	21%
Education[a]		
Low	35,996	25%
Medium	72,765	50%
High	36,448	25%
Duration of employment (years)		
<1	33,414	23%
1 ≤ 5	26,066	18%
5 ≤ 10	19,198	13%
10 ≤ 20	27,336	19%
20 ≤ 30	25,233	17%
30+	13,962	10%
Length of follow-up (years)		
Mean [Std]	30.1 [6.7]	
Range [Min, Max]	[0.02, 43.0]	

[a]SES was categorized as low, medium, or high based on the highest ever census block-level mean educational attainment.

Table 2. Number of workers and deaths by dose categories for ischemic heart disease and leukemia other than chronic lymphocytic leukemia.

Dose categories (mGy)	# Workers	# Deaths
Ischemic heart disease		
<5	30,152	1,876
5–24	53,861	2,179
25–49	27,439	914
50–74	12,369	395
75–99	6,913	180
100–124	4,542	134
125–149	2,880	73
150–199	3,388	91
200–249	1,699	41
250+	1,966	71
TOTAL	**145,209**	**5,954**
Leukemia		
<5	27,803	78
5–49	86,798	192
50–99	18,237	42
100+	12,371	25
TOTAL	**145,209**	**337**

a probabilistic scoring system, were used to obtain fact of death information. Both underlying and contributing causes of death were obtained via linkage with the National Death Index (NDI), state mortality data, or death certificates obtained from state departments of health. Two outcomes were chosen to investigate the effect of breaking the 'rare event' assumption of Poisson regression. The 'rare event' examined was death from leukemia other than chronic lymphocytic leukemia (hereafter called non-CLL), and for comparison, the 'prevalent event' outcome examined was death from ischemic heart disease (IHD). International classification of disease (ICD) codes used to define non-CLL and IHD are shown in Supplemental Table 2. For IHD, contributing causes were not included in the analyses.

Statistical modeling

For modeling purposes, to allow for a possible latent period between radiation exposure and any effect consequent to it, doses were lagged by two years for analyses of non-CLL deaths (i.e. doses were excluded if they occurred within two years before the date of death) and doses were lagged by 10 years for analyses of IHD deaths. Lagging was achieved by constructing an annual-level file for each cohort with each record representing one year of follow-up. Radiation exposure in each year was included in the annual-level file and then lagged two or ten years. For doses achieved prior to the follow-up window for each cohort, doses were summed to the first year of follow-up, less the lag period. Follow-up time was calculated from the start of follow-up to the first of death, loss to follow-up, age 95, or the end of the study period for each cohort. All cohorts had available measures of gender and year of birth which were included as covariates to all simple statistical models. Year of birth was categorized in five- or 10-year categories for each cohort.

Cox Proportional Hazards models were conducted using SAS/STAT software (SAS Institute 2013), R software (R Core Team 2012), and Epicure (Preston et al. 2015). In the Cox models, age was used as the timescale for the hazard function. Doses were treated as time-dependent exposures at the annual-level and were examined both in categories and as a linear (continuous) function. Hazard ratios (HR) and Wald's 95% confidence intervals (CI) were reported for each model using SAS and R (glm function within the stats package), while the Peanuts program of Epicure was utilized to produce ERR estimates and Wald's 95% CI. Four piece-wise exponential Poisson models using log-linear regression for the rate function were developed with varying cutpoints for age strata from very fine to broad categories using both R and the Amfit program in Epicure to produce ERR estimates. The relative risks (RR) and ERRs with corresponding 95% confidence intervals were compared to those from the Cox models. For the continuous models, the beta parameter (slope), standard error of the slope, and p-value for linear trend were also included for comparison.

Results

The demographic characteristics of the NPP cohort are presented in Table 1. The cohort is primarily male, mostly born after 1940, followed for an average of 30 years, have medium or low education levels, and were employed for less than 20 years. Table 2 presents the distribution of workers and deaths across dose categories. There were 5,954 total deaths due to IHD and 337 deaths due to non-CLL.

Comparison of models for leukemia other than chronic lymphocytic leukemia

Estimates of risk for non-CLL from the Cox proportional hazards model are compared with those from four Poisson models for categorical and continuous dose in Table 3. Between

models, estimates tended to agree within each category except for the Poisson model with the fewest age categories, where risk estimates were elevated compared to the other models. However, 95% CI were still overlapping, indicating there was no statistical difference. A similar trend was observed for the continuous estimates for RR and ERR at 100 mGy, such that the Cox and Poisson models with numerous age categories had very close agreement. The Poisson model with only three age categories had the highest risk estimates [RR 1.2 (95% CI 1.02, 1.44) and ERR 0.20 (95% CI 0.02. 0.38)] and also exhibited the only statistically significant result, with all other CIs including the null values of RR = 1 or ERR = 0. All three software systems evaluated (SAS, R, and Epicure) produced the exact same results for the Cox estimates (data not shown). Only R and Epicure were used to evaluate the Poisson models and they produced the exact same estimates for each of the models. Further, the ERR estimates produced in Epicure tended to agree very closely with the HR or RR estimates, following the theoretical relationship ERR = RR-1 (Preedy and Watson 2010). The results are not shown for brevity, but similar trends were observed in the other five cohorts for non-CLL estimates. However, the smaller cohorts (Mound and MCW) tended to exhibit slightly more pronounced differences in model estimates, with the Poisson models with fewer age categories tending to produce higher estimates compared to other models. Additionally, the ERRs did not track as closely to the RR-1, regardless of Cox or Poisson, for the smaller cohorts compared to the larger cohorts.

Comparison of models for ischemic heart disease

Estimates of risk for IHD from the Cox proportional hazards model is compared with those from four Poisson models for categorical and continuous dose in Table 4. As seen for non-CLL, estimates for IHD risk within categories tended to closely agree between the Cox and Poisson models with age cutpoints for every event and with small 5-year age categories. However, agreement tended to decrease as age categories widened. While several statistically significant category estimates were observed, the overall linear trend, both for RR and ERR, was null, indicating no significant dose-response for any model except for the Poisson model with the fewest age categories [RR 1.16 (95% CI 1.11, 1.21) and ERR 0.15 (95% CI 0.11, 0.19)]. Once again, the computer software programs had exact agreement and the ERR and RR values were comparable. Similar to the patterns observed for non-CLL, more variation in risk estimates were observed in the smaller cohorts compared to the large (data not shown).

Discussion

In this comparison of common statistical methods utilized for analysis of time-dependent dose response relationship for radiation in occupational cohorts, the results confirm that comparable estimates of risk (both RR and ERR) can be expected from Cox and Poisson models, regardless of software utilized, as long as appropriately narrow categories of age are utilized to control the confounding of age in Poisson models. Other theoretical and empirical comparisons of Cox and Poisson methods for time-dependent analyses have also reported instances of close agreement in estimates (Pearce et al. 1988; Callas et al. 1994, 1998; Steenland and Deddens 2004). In fact, Cox and Poisson estimates are expected to converge as the age strata in a Poisson model are made infinitely small (Checkoway et al. 2004). However, potential issues may arise when age is a strong confounder, which is often the case for occupational studies (Breslow et al. 1983),

Table 3. Comparison of different relative risk models for leukemia other than chronic lymphocytic leukemia for Nuclear Power Plant workers by radiation dose (mGy).

	Cox Proportional Hazards model		Poisson model with extremely fine age categories[a]		Poisson model with 5-year interval age categories[b]		Poisson model with 10-year interval age categories[c]		Poisson model with few age categories[d]	
					Categorical dose model					
Dose category	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
<5	1.0	Ref	1.0	Ref	1.0	Ref	1.0	Ref	1.0	Ref
5–49	1.15	0.88, 1.50	1.17	0.90, 1.54	1.15	0.88, 1.51	1.16	0.89, 1.52	1.29	0.99, 1.69
50–99	1.42	0.97, 2.09	1.45	0.99, 2.13	1.43	0.97, 2.09	1.44	0.98, 2.11	1.64	1.12, 2.40
100+	1.27	0.81, 2.01	1.3	0.82, 2.06	1.28	0.81, 2.02	1.29	0.82, 2.04	1.47	0.93, 2.33
					Continuous dose model					
	Beta [std] *p value*	RR 95% CI	Beta [std] *p-value*	RR 95% CI	Beta [std] *p-value*	RR 95% CI	Beta [std] *p-value*	RR 95% CI	Beta [std] *p-value*	RR 95% CI
RR at 100 mGy	0.00146 [0.00092] 0.11	1.16 0.97, 1.39	0.00149 [0.00092] 0.10500	1.16 0.97, 1.39	0.00146 [0.00092] 0.11200	1.16 [0.97,1.39]	0.00151 [0.0009] 0.10200	1.16 [0.97, 1.39]	0.00192 [0.0009] 0.03200	1.21 [1.02, 1.44]
ERR at 100 mGy	0.0015 [0.0015] 0.238	0.15 −0.04, 0.33	0.0016 [0.0014] 0.238	0.16 −0.05, 0.39	0.0019 [0.0015] 0.194	0.16 −0.03, 0.38	0.0022 [0.0015] 0.153	0.16 −0.03, 0.33	0.0026 [0.0008] 0.09	0.20 0.02, 0.38

Leukemia other than chronic lymphocytic leukemia cases based on underlying and contributing causes of death. Radiation dose has 2 year lag applied. All models adjusted for gender and year of birth (categories).
[a]Age cut points at all event points.
[b]Age cut points at 5-year intervals: <30, <35, <40, <45, <50, <55, <60, <65, <70, < 75, <80, <85, <90, 90+ years.
[c]Age cut points at 10-year intervals: <30, <40, <50, <60, <70, <80, <90, 90+ years.
[d]Age cut points at < 30, < 50, and 50+ year based on distribution.

Table 4. Comparison of different relative risk models for ischemic heart disease for Nuclear Power Plant workers by radiation dose (mGy).

	Cox Proportional Hazards model		Poisson model with extremely fine age categories[a]		Poisson model with 5-year interval age categories[b]		Poisson model with 10-year interval age categories[c]		Poisson model with few age categories[d]	
Categorical dose model										
Dose category	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
<5	1.0	Ref	1.0	Ref	1.0	Ref	1.0	Ref	1.0	Ref
5–24	1.23	1.15, 1.31	1.23	1.15, 1.32	1.24	1.16, 1.33	1.30	1.21, 1.38	1.86	1.75, 1.99
25–49	1.16	1.07, 1.26	1.16	1.07, 1.26	1.17	1.08, 1.28	1.23	1.13, 1.34	1.81	1.67, 1.97
50–74	1.20	1.07, 1.34	1.20	1.07, 1.34	1.21	1.08, 1.36	1.27	1.14, 1.42	1.89	1.69, 2.11
75–99	1.04	0.89, 1.21	1.04	0.89, 1.21	1.05	0.90, 1.22	1.10	0.94, 1.29	1.63	1.40, 1.90
100–124	1.27	1.06 1.52	1.27	1.06, 1.52	1.28	1.07, 1.53	1.35	1.13, 1.61	2.01	1.68, 2.39
125–149	1.08	0.86, 1.37	1.08	0.86, 1.37	1.10	0.87, 1.39	1.15	0.91, 1.46	1.70	1.35, 2.15
150–199	1.10	0.89, 1.36	1.10	0.89, 1.36	1.11	0.90 1.37	1.17	0.94, 1.44	1.73	1.40, 2.13
200–249	0.99	0.73, 1.36	0.99	0.73, 1.36	1.00	0.74, 1.37	1.05	0.77, 1.44	1.54	1.13, 2.09
250+	1.05	0.83, 1.33	1.05	0.83, 1.33	1.06	0.84, 1.35	1.12	0.88, 1.42	1.68	1.33, 2.14
Continuous dose model										
	Beta [std] *p-value*	RR 95% CI	Beta [std] *p-value*	RR 95% CI	Beta [std] *p-value*	RR 95% CI	Beta [std] *p-value*	RR 95% CI	Beta [std] *p-value*	RR 95% CI
RR at 100 mGy	−0.00004 [0.00024] 0.87	1.00 0.95, 1.05	−0.00004 [0.00024] 0.86	1.00 0.95, 1.04	−0.00000 [0.00024] 0.99	1.00 0.95, 1.05	0.00016 [0.0002] 0.50	1.02 0.97, 1.06	0.0015 [0.0002] < 0.001	1.16 1.11, 1.21
ERR at 100 mGy	−0.00004 [0.0003] >0.5	0.00 −0.05, 0.04	−0.00004 [0.0003] >0.5	0.00 −0.05, 0.04	0.00000 [0.00024] >0.5	0.00 −0.05, 0.05	0.0002 [0.0002] 0.442	0.02 −0.03, 0.07	0.0015 [0.0002] <0.001	0.15 0.11, 0.19

Ischemic heart disease based on underlying cause of death only. Radiation dose has 10 year lag applied. All models adjusted for gender and year of birth (categories).
[a]Age cut points at all event points.
[b]Age cut points at 5-year intervals: <30, <35, <40, <45, <50, <55, <60, <65, <70, < 75, <80, <85, <90, 90+ years.
[c]Age cut points at 10-year intervals: <30, <40, <50, <60, <70, <80, <90, 90+ years.
[d]Age cut points at < 30, < 50, and 50+ year based on distribution.

and strata for Poisson regression cannot be made infinitely small, typically because there are too few deaths to support numerous categories. One study (Callas et al. 1998) evaluated numerous models with varying widths of age and calendar time categories, concluding that when wide age intervals were used, there was a marked change with as much as 8% difference in Poisson estimates compared to those from Cox. The intervals for calendar time had less of a pronounced impact. They concluded that Poisson and Cox proportional hazards yield nearly identical estimates of risk and confidence intervals except when confounding by age could not be closely controlled (Callas et al. 1998). It is important to obtain maximum possible control for age in mortality studies as death rates are expected to rise with increasing age (Breslow et al. 1983; Korn et al. 1997). Since the Cox proportional hazards model produces a stratum for each case, age-strata are not explicitly included as terms in the model. Thus, age is the 'time' variable that defines the risk set at the time (age) for each case such that the rate ratio is independent of age (Breslow et al. 1983; Stayner et al. 1995; Checkoway et al. 2004; Allison 2010). This provides an additional advantage of the Cox proportional hazards model over the Poisson by assuming that the rate is constant within strata (Checkoway et al. 2004). The stricter assumption of the Poisson model of constant hazard relative to the Cox model's proportional hazards one, could lead to bias in estimates if time intervals are wide (Callas et al. 1998; Richardson and Loomis 2004; Allison 2010), as was consistently seen in this evaluation for the Poisson models with fewer age strata.

Nearly the exact same estimates and CIs were observed for the Cox model and the Poisson model with age strata for every event for all outcome scenarios across all cohorts. Others have described and evaluated an 'ungrouped' Poisson methodology to avoid potential biases associated with stratification (Loomis et al. 2005). In the ungrouped analyses, the analytic dataset was constructed in the same manner as for Cox proportional hazards such that there is an unique observation for each unit of person time at risk, resulting in a file with one observation for each year of follow up for each person (in contrast to the tabular form for stratified Poisson analyses). The authors of that study concluded that the ungrouped Poisson analysis can be used to estimate quantitative exposure-response relationships and address concerns regarding potential bias in categorical analyses in the same manner as Cox proportional hazards.

It was reassuring to observe that the choice of statistical software package had no impact on risk estimates for models specified in the same manner. It was expected that the estimates of ERR and HR (Cox) or RR (Poisson) would mostly adhere to the theoretical expectation of ERR = RR-1 (Preedy and Watson 2010), though some deviation was observed for the smaller cohorts, but not so important as to affect the interpretation with regards to significance.

In conclusion, as computational power is no longer the burden today as it has been in the past, the results of this evaluation support the use of the Cox proportional hazards or the ungrouped Poisson approach to analyzing time-dependent dose–response relationships to ensure that maximum control of confounding by age is achieved in studies of mortality for cohorts occupationally exposed to radiation. Stratified Poisson models may be limited in their ability to precisely control for age effects, particularly when the cohort

and/or number of deaths are small. As one of the goals of the study of a *Million U.S. Radiation Workers and Atomic Veterans* (MWS) is to promote harmonization of many different cohorts utilizing the most optimal methods for vital status tracing, dosimetry, and statistical analyses, it was important to evaluate methodological inconsistencies and uncertainties observed in the existing literature. This methods paper highlights the importance of careful consideration of the statistical approach and demonstrates both the prudence in performing sensitivity analyses and the need for authors to be explicit in the description of statistical approaches in the literature to allow for valid comparisons between studies. There are several other aspects of statistical modelling approaches that were beyond the scope of this current evaluation, including: incorporation of dosimetric errors, evaluation of fit with regards to dose–response shapes other than linear, model implications for excess absolute risk models, and the inclusion or exclusion of important covariates, both fixed and time-varying (such as duration of employment). Additional efforts are underway within the MWS to optimize the approach to addressing these issues in order to refine and enhance estimates of risk to improve our understanding of low-dose radiation exposures.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding

This work was supported by the Oak Ridge Associated Universities Directed Research and Development program, an internally funded competitive grant program for ORAU employees and member universities. Additionally, the data analyzed were provided through support by a grant from the U.S. Nuclear Regulatory Commission (NRC-HQ-60-14-G-0011), a grant from the Centers for Disease Control and Prevention (5UE1EH000989), a grant from the National Aeronautics and Space Administration (Grant No. NNX15AU88G), a grant from the National Cancer Institute (Grant No. U01 CA137026), a Discovery Grant from the Vanderbilt-Ingram Cancer Center (Center no. 404-357-9682), and grants from the U.S. Department of Energy (Grant No. DE-SC0008944 and Grant No. DE-AU0000042 awarded to the National Council on Radiation Protection and Measurements, which included interagency support from the U.S. Nuclear Regulatory Commission, the U.S. Environmental Protection Agency and the National Aeronautics and Space Administration).

Notes on contributors

Ashley P. Golden is a biostatistician and project manager at Oak Ridge Associated Univerities were she directs and supports multi-disciplinary projects in the areas of occupational epidemiology, radiation exposure and dosimetry, medical surveillance, survey analyses, and environmental assessments. She has been a collaborator on the Million Person Study of Low-Dose Health Effects for nearly six years.

Sarah S. Cohen is a Principal Epidemiologist at EpidStat Institute where she directs observational research studies in the areas of pharmacoepidemiology, nutritional epidemiology, and occupational epidemiology as well as leads large data management projects and statistical analyses. She is also an Adjunct Assistant Research Professor of Medicine in the Department of Medicine at Vanderbilt University School of Medicine. She has been a collaborator on the Million Person Study of Low-Dose Health Effects for nearly twenty years, providing analytic support as well as co-authoring numerous publications.

Heidi Chen is a Research Assistant Professor of Biostatistics in the Vanderbilt Center for Quantitative Sciences. Her Ph.D. is in biomathematics from North Carolina State University and her research focuses on applying mathematical, statistical and computational methods to fit models for biological systems. Since joining the Department of Biostatistics, Vanderbilt University Medical Center in 2003, her research has involved identifying biomarkers and building models to predict disease progression and clinical outcome for cancer patients. Heidi has been a statistical collaborator on the Million Person Study of Low-Dose Health Effects since 2013.

Elizabeth D. Ellis is an occupational epidemiologist at Oak Ridge Associated Universities. She has been studying the health effects of chronic low dose ionizing radiation on the Department of Energy nuclear workers for over 35 years and has been a collaborator on the Million Person Study of Low-Dose Health Effects for over 10 years. She is a member of an International Commission of Radiation Protection Task Group reviewing the health effects of alpha emitters. She is also active in protection of human participants in research serving on several Institutional Review Boards.

John D. Boice, Jr. is President of the National Council on Radiation Protection and Measurements and Professor of Medicine at Vanderbilt University. He is an international authority on radiation effects and served on the Main Commission of the International Commission on Radiological Protection and on the United Nations Scientific Committee on the Effects of Atomic Radiation. He directs the Million Person Study of Low-Dose Health Effects.

ORCID

Ashley P. Golden http://orcid.org/0000-0002-9430-7702
Elizabeth D. Ellis http://orcid.org/0000-0003-2871-1587

References

Allison P. 2010. Survival analysis using the SAS system: a practical guide. 2nd ed. Cary, NC: SAS Institue, Inc.

Anderson JL, Daniels RD, Fleming DA, Tseng CY. 2012. Exposure assessment for a cohort of workers at a former uranium processing facility. J Expo Sci Environ Epidemiol. 22:324–330.

Boice JD, Cohen SS, Mumma MT, Chadda B, Blot WJ. 2007. Mortality among residents of Uravan, Colorado who lived near a uranium mill, 1936–84. J Radiol Prot. 27:299–319.

Boice JD, Cohen SS, Mumma MT, Chadda B, Blot WJ. 2008. A cohort study of uranium millers and miners of Grants, New Mexico, 1979–2005. J Radiol Prot. 28:303–325.

Boice JD, Cohen SS, Mumma MT, Dupree-Ellis E, Eckerman KF, Leggett RW, Boecker B, Brill A, Henderson B. 2006. Mortality among radiation workers at Rocketdyne (Atomics International), 1948–1999. Radiat Res. 166:98–115.

Boice JD, Cohen SS, Mumma MT, Ellis ED, Cragle DL, Eckerman KF, Wallace PW, Chadda B, Sonderman JS, Wiggs LD, et al. 2014. Mortality among mound workers exposed to Polonium-210 and other sources of radiation, 1944–1979. Radiat Res. 181:208–228.

Boice JD, Cohen SS, Mumma MT, Ellis ED, Eckerman KF, Leggett RW, Boecker BB, Brill AB, Henderson BE. 2011. Updated mortality analysis of radiation workers at Rocketdyne (Atomics International), 1948–2008. Radiat Res. 176:244–258.

Boice JD, Cohen SS, Mumma MT, Golden AP, Hagemeyer D, Yoder C, Dauer LT. In press. Leukemia Among Industrial Radiographers. Int J Rad Biol.

Boice JD, Cohen SS, Mumma MT, Hagemeyer D, Chen H, Golden AP, Yoder RC, Dauer LT. In press. Leukemia among Nuclearl Power Plant Workers Monitored 1957-1984 in the United States. Int J Rad Biol.

Boice JD, Ellis ED, Golden AP, Girardi DJ, Cohen SS, Chen H, Mumma MT, Shore RE, Leggett RW. 2018. The past informs the future: an overview of the million worker study and the mallinckrodt chemical works cohort. Health Phys. 114:381–385.

Boice JD, Till J, Mumma MT, Chen H, Cohen SS. In press. Mortality among US military participants at seven aboveground nuclear weapons test series. Int J Rad Biol.

Bouville A, Toohey RE, Boice JD, Beck HL, Dauer LT, Eckerman KF, Hagemeyer D, Leggett RW, Mumma MT, Napier B, et al. 2015. Dose Reconstruction for the Million Worker Study: Status and Guidelines. Health Phys. 108:206–220.

Breslow N. 1974. Covariance analysis of censored survival data. Biometrics. 30:89–89.

Breslow N, Day N. 1987. The Design and Analysis of Cohort Studies. Lyon, France: International Agency fo Research on Cancer.

Breslow NE, Lubin J, Marek P, Langholz D. 1983. Multiplicative models and cohort analysis. J Am Stat Assoc. 78:1–12.

Callas PW, Pastides H, Hosmer DW. 1994. Survey of methods and statistical models used in the analysis of occupational cohort studies. Occup Environ Med. 51:649–655.

Callas PW, Pastides H, Hosmer DW. 1998. Empirical comparisons of proportional hazards, Poisson, and logistic regression modeling of occupational cohort data. Am J Ind Med. 33:33–47.

Cardis E, Vrijheid M, Blettner M, Gilbert E, Hakama M, Hill C, Howe G, Kaldor J, Muirhead CR, Schubauer-Berigan M, et al. 2007. The 15-country collaborative study of cancer risk among radiation workers in the nuclear industry: Estimates of radiation-related cancer risks. Radiat Res. 167:396–416.

Checkoway H, Pearce N, Kreibel D. 2004. Research methods in occupational epidemiology. 2nd ed. New York, New York: Oxford University Press.

Cox DR. 1972. Regression models and life-tables. J R Stat Soc Ser B. 34: 187–220.

Cox DR, Oakes D. 1984. Analysis of Survival DAta. New York: Chapman and Hall.

Dauer LT, Bouville A, Toohey RE, Boice JD, Beck H, Eckerman KF, Hagemeyer D, Leggett RW, Mumma MT, Napier B, et al. 2018. Dosimetry and uncertainty approaches for the million person study of low-dose radiation health effects: overview of the recommendations in NCRP Report No. 178. Int J Rad Biol. Nov 19:1–10. Epub 2018 Nov 19.

Dupree-Ellis E, Watkins J, Ingle J, Phillips J. 2000. External radiation exposure and mortality in a cohort of Uranium processing workers. Am J Epidemiol. 152:91–95.

Eidemuller M, Jacob P, Lane R, Frost S, Zablotska L. 2012. Lung cancer mortality (1950-1999) among Eldorado uranium workers: a comparison of models of carcinogenesis and empirical excess risk models. PLoS One. 7:e41431.

Ellis ED, Boice JD, Golden AP, Girardi DJ, Cohen SS, Mumma MT, Shore RE, Leggett RW, Kerr GD. 2018. Dosimetry is key to good epidemiology: workers at mallinckrodt chemical works had seven different source exposures. Health Phys. 114:386–397.

Frome EL. 1983. The analysis of rates using Poisson regression models. Biometrics. 39:665–674.

Frome EL, Checkoway H. 1985. Epidemiologic programs for computers and calculators. Use of Poisson regression models in estimating incidence rates and ratios. Am J Epidemiol. 121:309–323.

Golden AP, Ellis ED, Cohen SS, Mumma MT, Leggett RW, Wallace PW, Watkins J, Shore RE, Boice JD. In press. Updated mortality analysis of the Mallinckrodt uranium processing workers, 1942-2012. Int J Rad Biol.

Guseva Canu I, Cardis E, Metz-Flamant C, Caër-Lorho S, Auriol B, Wild P, Laurier D, Tirmarche M. 2010. French cohort of the uranium processing workers: mortality pattern after 30-year follow-up. Int Arch Occup Environ Health. 83:301–308.

Guseva Canu I, Garsi J, Caër-Lorho S, Jacob S, Collomb P, Acker A, Laurier D. 2012. Does uranium induce circulatory diseases? First results from a French cohort of uranium workers. Occup Environ Med. 69:404–409.

Hornung RW, Pinney SM, Lodwick J, Killough GG, Brewer DE, Nasuta J. 2008. Estimation of radon exposures to workers at the Fernald Feed Materials Production Center 1952-1988. J Expo Sci Environ Epidemiol. 18:512–523.

Hsu W-L, Preston DL, Soda M, Sugiyama H, Funamoto S, Kodama K, Kimura A, Kamada N, Dohy H, Tomonaga M, et al. 2013. The incidence of leukemia, lymphoma and multiple myeloma among atomic bomb survivors: 1950-2001. Radiat Res. 179:361–382.

Ingram DD, Kleinman JC. 1989. Empirical comparisons of proportional hazards and logistic regression models. Stat Med. 8:525–538.

Korn EL, Graubard BI, Midthune D. 1997. Time-to-event analysis of longitudinal follow-up of a survey: Choice of the time-scale. Am J Epidemiol. 145:72–80.

Kreuzer M, Dufey F, Laurier D, Nowak D, Marsh JW, Schnelzer M, Sogl M, Walsh L. 2015. Mortality from internal and external radiation exposure in a cohort of male German uranium millers, 1946–2008. Int Arch Occup Environ Health. 88:431–441.

Loomis D, Richardson DB, Elliott L. 2005. Poisson regression analysis of ungrouped data. Occup Environ Med. 62:325–329.

McGeoghegan D, Binks K. 2006. The mortality and cancer morbidity experience of workers at the Springfields uranium production facility, 1946–1995 (vol 20, pg 111, 2006) [Correction]. J Radiol Prot. 26: 455–455.

Mumma MT, Cohen SS, Ellis ED, Boice JD. 2018. Obtaining vital status and cause of death on a million persons. Int J Rad Biol. doi: 10.1080/09553002.2018.1539884. [Epub ahead of print].

Ozasa K, Shimizu Y, Suyama A, Kasagi F, Soda M, Grant EJ, Sakata R, Sugiyama H, Kodama K. 2012. Studies of the mortality of atomic bomb survivors, report 14, 1950-2003: an overview of cancer and noncancer diseases. Radiat Res. 177:229–243.

Ozasa K, Takahashi I, Grant EJ, Kodama K. 2017. Cardiovascular disease among atomic bomb survivors. Int J Radiat Biol. 93:1145–1150.

Pearce N, Checkoway H, Dement J. 1988. Exponential models for analyses of time-related factors, illustrated with asbestos textile worker mortality data. J Occcup Med. 30:517–522.

Pinkerton LE, Bloom TF, Hein MJ, Ward EM. 2004. Mortality among a cohort of uranium mill workers: an update. Occup Environ Med. 61: 57–64.

Preedy VR, Watson RR. 2010. Excess Relative Risk. Handbook of Disease Burdens and Quality of Life Measures. New York, NY: Springer.

Preston DL, Lubin J, Pierce D, McConney M, Shilnikova N. 2015. Epicure Risk Regression and Person-Year Computation Software: Command summary and user guide Ottawa, Canada: Risk Sciences International.

Preston DL, Ron E, Tokuoka S, Funamoto S, Nishi N, Soda M, Mabuchi K, Kodama K. 2007. Solid cancer incidence in atomic bomb survivors: 1958-1998. Radiat Res. 168:1–64.

R Core Team 2012. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Richardson DB, Cardis E, Daniels RD, Gillies M, Haylock R, Leuraud K, Laurier D, Moissonnier M, Schubauer-Berigan MK, Thierry-Chef I, et al. 2018. Site-specific solid cancer mortality after exposure to ionizing radiation: a cohort study of workers (INWORKS). Epidemiology. 29: 31–40.

Richardson DB, Loomis D. 2004. The impact of exposure categorisation for grouped analyses of cohort data. Occup Environ Med. 61:930–935.

SAS Institute I 2013. SAS/STAT software Version 9.4 for Windows. Cary, NC: The SAS System.

Schubauer-Berigan M, Daniels R, Pinkerton L. 2009. Radon Exposure and Mortality Among White and American Indian Uranium miners: an update of the Colorado Plateau cohort. Am J Epidemiol. 169:718–730.

Sera N, Hida A, Imaizumi M, Nakashima E, Akahoshi M. 2013. The association between chronic kidney disease and cardiovascular disease risk factors in atomic bomb survivors. Radiat Res. 179:46–52.

Shimizu Y, Kodama K, Nishi N, Kasagi F, Suyama A, Soda M, Grant EJ, Sugiyama H, Sakata R, Moriwaki H, et al. 2010. Radiation exposure and circulatory disease risk: Hiroshima and Nagasaki atomic bomb survivor data, 1950-2003. Br Med J. 340:8.

Shore RE, Beck HL, Caffrey EA, Davis S, Grogan HA, Mettler FA, Preston RJ, Till JE, Wakeford R, Walsh L, et al. 2018. Implications of recent epidemiologic studies for the linear nonthreshold model and radiation protection. J Rad Prot. 38(3). DOI: 10.1088/1361-6498/aad348.

Silver SR, Bertke SJ, Hein MJ, Daniels RD, Fleming DA, Anderson JL, Pinney SM, Hornung RW, Tseng CY. 2013. Mortality and ionising radiation exposures among workers employed at the Fernald Feed Materials Production Center (1951-1985). Occup Environ Med. 70: 453–463.

Stayner L, Smith R, Bailer J, Luebeck EG, Moolgavkar SH. 1995. Modeling Epidemiologic Studies of Occupational Cohorts for the Quantitative Assessment of Carcinogenic Hazards. Am J Ind Med. 27:155–170.

Steenland K, Deddens JA. 2004. A practical guide to dose-response analyses and risk assessment in occupational epidemiology. Epidemiology. 15:63–70.

Yamada M, Kasagi F, Mimori Y, Miyachi T, Ohshita T, Sasaki H. 2009. Incidence of dementia among atomic-bomb survivors – Radiation Effects Research Foundation Adult Health Study. J Neurol Sci. 281: 11–14.

Yiin JH, Anderson JL, Bertke SJ, Tollerud DJ. 2018. Dose-response relationships between internally-deposited uranium and select health outcomes in gaseous diffusion plant workers, 1948-2011. Am J Ind Med. 61:605–614.

Yiin JH, Anderson JL, Daniels RD, Bertke SJ, Fleming DA, Tollerud DJ, Tseng CY, Chen PH, Waters KM. 2017. Mortality in a combined cohort of uranium enrichment workers. Am J Ind Med. 60:96–108.

Yiin JH, Anderson JL, Daniels RD, Seel EA, Fleming DA, Waters KM, Chen PH. 2009. A nested case-control study of multiple myeloma risk and uranium exposure among workers at the oak ridge gaseous diffusion plant. Radiat Res. 171:637–645.

Zablotska LB, Fenske N, Schnelzer M, Zhivin S, Laurier D, Kreuzer M. 2018. Analysis of mortality in a pooled cohort of Canadian and German uranium processing workers with no mining experience. Int Arch Occup Environ Health. 91:91–103.

Zablotska LB, Lane RS, Frost SE. 2013. Mortality (1950–1999) and cancer incidence (1969–1999) of workers in the Port Hope cohort study exposed to a unique combination of radium, uranium and γ-ray doses. BMJ Open. 3:e002159.

Zablotska LB, Lane RS, Thompson PA. 2014. A reanalysis of cancer mortality in Canadian nuclear workers (1956-1994) based on revised exposure and cohort data. Br J Cancer. 110:214–223.

Zhivin S, Canu I, Samson E, Laurent O, Grellier J, Collomb P, Zablotska L, Laurier D. 2016. Mortality (1968-2008) in a french cohort of uranium enrichment workers potentially exposed to rapidly soluble uranium compounds. Occup Environ Med. 73:167–174.