Occupational Cohort Time Scales

David C. Deubner, MD, MPH and H. Daniel Roth, PhD

Purpose: This study explores how highly correlated time variables (occupational cohort time scales) contribute to confounding and ambiguity of interpretation. **Methods:** Occupational cohort time scales were identified and organized through simple equations of three time scales (relational triads) and the connections between these triads (time scale web). The behavior of the time scales was examined when constraints were imposed on variable ranges and interrelationships. **Results:** Constraints on a time scale in a triad create high correlations between the other two time scales. These correlations combine with the connections between relational triads to produce association paths. High correlation between time scales leads to ambiguity of interpretation. **Conclusions:** Understanding the properties of occupational cohort time scales, their relational triads, and the time scale web is helpful in understanding the origins of otherwise obscure confounding bias and ambiguity of interpretation.

Time is intrinsic to epidemiology in basic concepts such as incidence and prevalence, as a denominator for time-based rates and as a surrogate for unmeasured or uncharacterized factors. Occupational exposures are often chronic, and therefore their measurement often involves some measure of time. The risk of disease often changes with calendar time, and even more so with respect to the age of an individual. The interrelation of these various time scales (Table 1, Fig. 1) in occupational epidemiology may be complex and offers various design and analytic choices to epidemiologic researchers. Readers of occupational epidemiology studies face the challenge of understanding whether the choices made by the researcher contribute to valid comparisons, which mitigate potential time scale–related confounding, or have the opposite effect, of introducing confounding strong enough to invalidate results. Thus, time-related confounding is a crucial concern, yet authors frequently do not explain the rationale for choosing one or more time scales for adjustment or risk set selection.

The objective of this paper was to report our current understanding of the structure and behavior of occupational cohort time scales. It is organized as a didactic explanation of occupational cohort time scales, how they are organized into simple relational triads of the form A = B + C, such as *date of hire* = *date of birth* + *age at hire*, and the interconnection between these triads forming a time scale web. This explanation is followed by illustration of how relational triads behave when subjected to constraints on one or two time scales constituting the triad, how lagging exposure creates associations between time scales and lagged exposure, and how log transformation of exposure variables may influence these associations. This information is then applied in the form of a case example utilizing the studies cited above. We conclude with consideration of how researchers can use the underlying structure of occupational cohort times scales in design of analyses and how readers of occupational cohort epidemiology studies can evaluate the handling of time scales.

The impetus for this paper began with the paper by Sanderson et al[1], a cohort-nested case–control study of beryllium occupational exposure and lung cancer mortality. Lagging exposure changed odds ratios so profoundly that an unusual bias was suspected.[2] A subsequent paper by Schubauer-Berigan et al[3] implicitly acknowledged bias in Sanderson et al[1] by reporting that adjustment for *date of birth* changed the results. *Date of birth* was shown to be associated with lung cancer rates over the range of *date of birth* of the cohort. Adjustment with *age at hire* changed the results to a similar degree, a finding that was explained by the high correlation ($r = 0.91$) between *date of birth* and *age at hire*. A simulation study by Rothman and Mosquin[4] also identified *date of birth* as an important confounder in Sanderson et al.[1] Nevertheless, neither paper identified an association path linking either *date of birth* or *age at hire* to both lung cancer and exposure. And, an additional question was added—why was *date of birth* so highly correlated with *age at hire* in this cohort? The result was that we continued to seek fuller understanding of the structure and behavior of occupational cohort time scales, an understanding that we hoped would answer the question of the time scale association path in Sanderson et al[1] that was responsible for the confounding.

METHODS

We drew a diagram of the typical life course of subjects in an occupational cohort (Fig. 1) and identified 10 occupational cohort time scales (Table 1). Knowing some of these 10 time scales was linked by equations of the form A = B + C, we compiled a complete list of these equations (Table 1). We termed each of these 10 equations a "relational triad." Observing that each time scale is incorporated into three relational triads, we diagramed this structure to demonstrate these connections (Fig. 2). We termed this structure the "time scale web".

We explored how constraints on the ranges of *date of hire* and *age at hire* influenced the correlation between *date of birth* and *age at hire* by constructing a synthetic cohort with an arbitrary range of *date of hire* of 100 years, 1901 to 2000. In each of these years, we "hired" 70 subjects, one each with a hire age from 15 to 84 years, thus ensuring that in the overall data there was zero correlation between *date of hire* and *age at hire*. Because subjects had been given year of hire and age at hire only in whole years, a random number between 0 and 1 was added to each to specify a date and age within that year. For example, a subject assigned *date of hire* 1990 and *age at hire* 43 years might be further specified as 1990.64921 and 43.22577 by the addition of the random numbers. We calculated *date of birth* as *date of hire* − *age at hire*, and then calculated the correlation between *date of birth* and *age at hire* over several ranges of *date of hire* and three ranges of *age at hire* (Fig. 3). We also used an occupational cohort[5] database supplied by the National Institute for Occupational Safety and Health (Cincinnati, OH) under a data use agreement to ascertain the range of *date of hire* and *age at hire* for several beryllium materials manufacturing facilities and calculated for each the correlation between *date of birth* and *age at hire*. We applied the facility ranges of *date of hire* and *age at hire* as constraints on the synthetic cohort and calculated for each combination the correlation

From Materion Brush Inc. (Dr Deubner), Elmore; and Roth Associates (Dr Roth), Englewood, Ohio.
Funding for this project was provided by Materion Inc. Materion manufactures beryllium-containing materials.
The authors declare no conflicts of interest.

Address correspondence to: David C. Deubner, MD, MPH, VP Occupational and Environmental Medicine, Materion Inc., 14710 West Portage River S Rd, Elmore, OH 43416 (david.deubner@materion.com).

DOI: 10.1097/JOM.0000000000000412

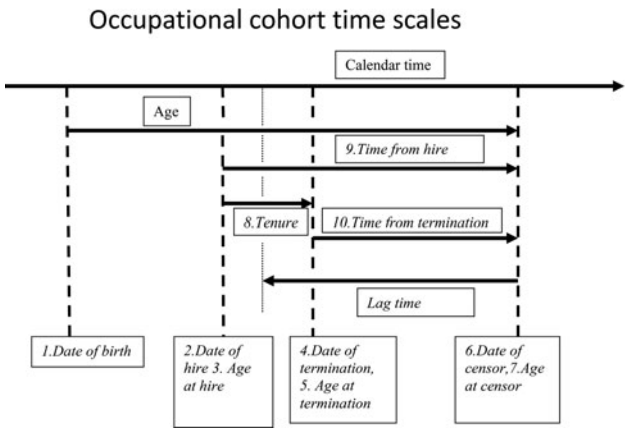

FIGURE 1. Time scales in occupational cohorts

TABLE 1. Occupational Cohort Time Scales and Relational Triads

	Occupational Cohort Time Scales
1	Date of birth
2	Date of hire
3	Age at hire
4	Date of termination
5	Age at termination
6	Date of censor
7	Age at censor
8	Tenure
9	Time from hire
10	Time from termination
	Relational Triads (Time Scale Equations)
1	Date of hire = date of birth + age at hire
2	Date of termination = date of birth + age at termination
3	Date of termination = date of hire + tenure
4	Age at termination = age at hire + tenure
5	Time from hire = tenure + time from termination
6	Date of censor = date of birth + age at censor
7	Date of censor = date of hire + time from hire
8	Date of censor = date of termination + time from termination
9	Age at censor = age at hire + time from hire
10	Age at censor = age at termination + time from termination

of *date of birth* with *age at hire*. We then compared the correlations derived from the synthetic cohort with the actual correlations (Table 2) as a demonstration of how constraints on time scales in relational triads influence the correlation between the time scales.

We prepared a diagram to illustrate how tenure, average exposure, and cumulative exposure change with time, and how this process is reversed by progressive lagging of exposure (increasing *lag time*) (Fig. 4). To illustrate how lagging exposure creates an association between *time from hire* and lagged exposure, we calculated in the controls in the Sanderson et al[1] study the correlation between lagged cumulative exposure and *time from hire* for lag 0, 10, and 20 years (Table 3). To assist with the case example we constructed a causal diagram to represent the progression of thinking (Fig. 5) and calculated the effect of adjustment with quartiles of several time scale variables on the odds ratios for quartiles of cumulative exposure lagged 10 years (Table 4).

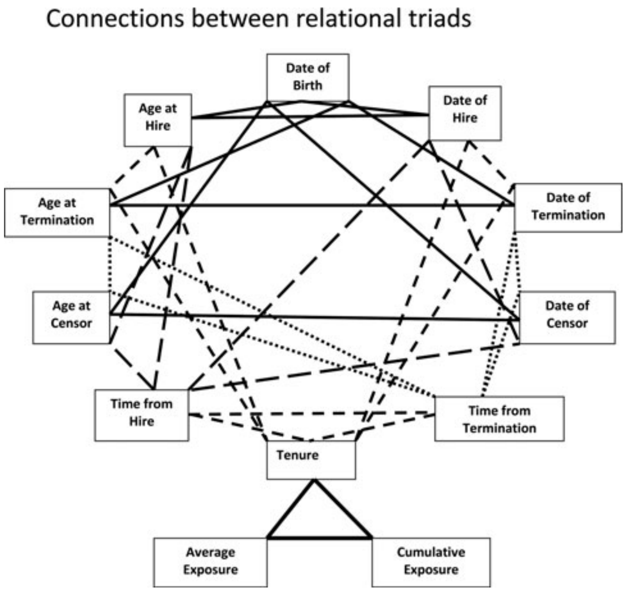

FIGURE 2. The time scale web: interrelationships between occupational cohort time scale relational triads plus their connection via tenure to the exposure variable relational triad.

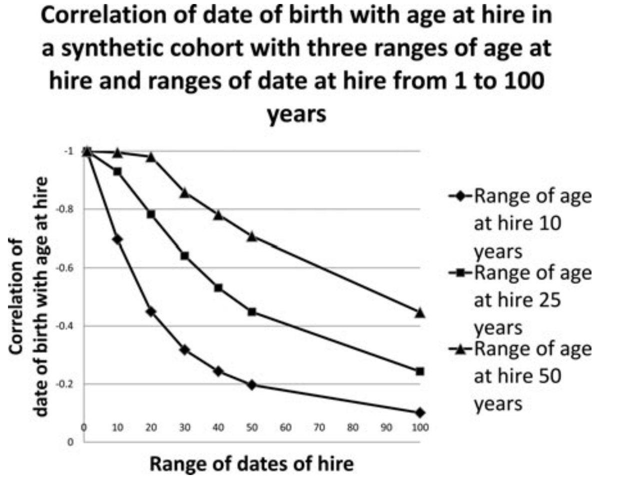

FIGURE 3. Example of how constraints on time scale variables in relational triads affects correlations: correlation in a synthetic cohort of *date of birth* with *age at hire* for different ranges of *date of hire* and *age at hire*.

RESULTS

Organization and Structure of Occupational Cohort Time Scales

From Fig. 1 we identified 10 occupational cohort time scales (Table 1). We were able to identify 10 relational triads with each time scale occurring in three relational triads (Table 1). The connections between the relational triads by their common time scales form the "time scale web," an arrangement of which is diagrammed in Fig. 2. As each time scale in a relational triad participates in two other relational triads, each relational triad can be connected via the corner timescales to six other triads. Figure 2 also depicts the connection

TABLE 2. Comparison of the Correlations of *Date of Hire* With *Age at Hire* in Eight Beryllium Facility Cohorts and All combined[8] With the Respective Correlations Derived From the Synthetic Cohort Applying the Same Ranges of *Date of Hire* and *Age at Hire*

Facility	Range of *Date of Hire* in Years	Range of *Age at Hire* in Years	Actual Cohorts Correlation *Date of Birth* With *Age at Hire* (Pearson *r*)	Synthetic Cohort Correlation *Date of Birth* With *Age at Hire* (Pearson *r*)
Luckey	12.1	47.6	− 0.98	− 0.97
Hazleton	12.2	40.1	− 0.93	− 0.96
Elmore	16.8	45.5	− 0.92	− 0.94
Lorain	17.6	64.6	− 0.98	− 0.96
Unknown	29.5	53.2	− 0.83	− 0.88
Multiple	37.5	43.9	− 0.81	− 0.76
Cleveland	37.6	53.3	− 0.82	− 0.82
Reading	40.7	61.2	− 0.85	− 0.83
All	40.7	65.3	− 0.81	− 0.85

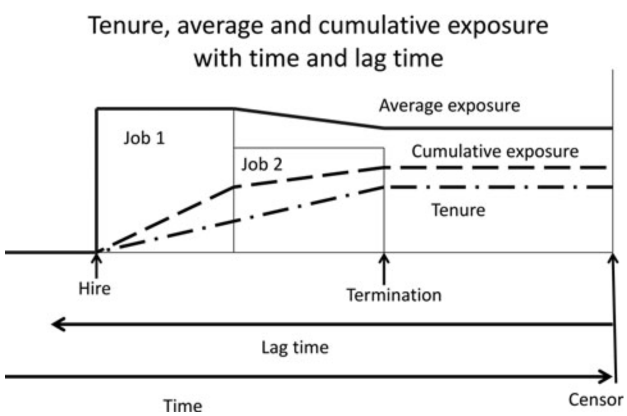

FIGURE 4. How lagging affects exposure metrics

via *tenure* of the time scales to the multiplicative exposure relational triad, *cumulative exposure* = *tenure* × *average exposure*.

Behavior of Relational Triads: Constraints

If one time scale in a relational triad is constrained to a narrow range, the other two become highly correlated. For example, in the relational triad *date of hire* = *date of birth* + *age at hire*, if the range of *age at hire* is narrow, as it might be in military recruits, the correlation between *date of hire* and *date of birth* would be high. To illustrate this we created a synthetic cohort and calculated the correlation between *date of birth* and *age at hire* for various combinations of range of *date of hire* and *age at hire*, demonstrating that the correlation increases with decreasing range of *date of hire* and with increasing range of *age at hire* (Fig. 3). That this is primarily influenced by constraint of the ranges of the time scales rather than by the detail of the respective distributions is illustrated by the agreement between correlations from real[5] and synthetic cohorts (Table 2).

Truncation of Time Scale Variables

In analyses in which controls are selected by risk set sampling, it is usual practice to truncate the experience of controls as of the time scale value for the case.[6] For example, if the time scale for risk set selection is *age at censor* of the case, information for controls matched to each case will be considered only up to the *age at censor* of the case. In the case example below, if the *age at censor* of the case

TABLE 3. Rank and Product Moment Correlations Between Truncated Time From Hire and Exposure, Unlagged, Lagged 10 Years and Lagged 20 Years. Product Moment Correlation was Calculated for Both the Original Data and the Natural Logarithm of the Original Data (After Substituting 0.1 for Zero in "Lagged Out" Subjects). The Subjects Are the 710 Controls in the Sanderson et al[1] Study*

	Unlagged	Lagged 10 Yrs	Lagged 20 Yrs
Tenure (employment duration)			
Rank	− 0.03	0.27	0.61
PM	0.14	0.26	0.36
PM Ln	− 0.01	0.45	0.72
Cumulative exposure			
Rank	0.01	0.29	0.61
PM	0.09	0.12	0.17
PM Ln	0.03	0.52	0.76
Average exposure			
Rank	0.05	0.28	0.59
PM	0.04	0.05	0.07
PM Ln	0.08	0.49	0.73
Maximum exposure			
Rank	0.09	0.32	0.60
PM	0.05	0.06	0.09
PM Ln	0.11	0.50	0.73

*Rank = rank order correlation (Spearman ρ = Pearson *r* of the ranks); PM = product moment correlation (Pearson *r*); Ln = natural logarithm of data after 0.1 substituted for zero.

were 65, with truncation the truncated *age at censor* for the case's five controls would also be 65 years. Therefore, in each case–control set, *age at censor* is set equal for the case and all controls, so the correlation of the other time scales in each of the three relational triads containing *age at censor* becomes 1.0 or −1.0. Thus, the correlation in each case–control set between *date of birth* and *date of censor* is 1.0, between *age at hire* and *time from hire* is −1.0, and between *age at termination* and *time from termination* is −1.0

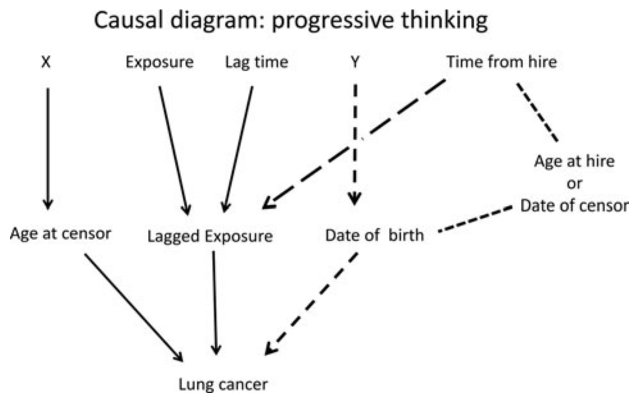

Figure 5. This figure, from left to right, depicts in causal diagram format, the progression of thinking regarding possible confounders in the study of Sanderson et al[1]. The first two paths (solid lines) were contemplated by Sanderson et al[1], the third starting with Y (fine dotted lines) was added by Schubauer-Berigan et al[6], and the path on the right (coarse segmented lines) by this study.

Lagging Exposure

Lagging exposure is a form of further truncation of exposure. Lagging is performed under the assumption that exposure occurring during the period between the induction[7] and the detection of disease does not contribute to the observed disease and that counting this exposure introduces error through misclassification of exposure. As *lag time* is progressively increased (Fig. 4) no change in exposure occurs until *lag time* equals *time from termination*. At this point tenure decreases linearly with increasing *lag time*, cumulative exposure also decreases, although not necessarily linearly, and average exposure may or may not change. When *lag time* = *time from hire*, all exposure metrics become zero. For a given *lag time*, exposure profiles with greater *time from termination* will have less truncation of tenure and cumulative exposure. Profiles with *time from hire* greater than the *lag time* will avoid all exposure lagging to zero. The result is that, with lagging, the correlations between both *time from termination* and lagged exposure and *time from hire* and lagged exposure move in a positive direction. This phenomenon is illustrated by comparing the correlation between *time from hire* and exposure lagged 0, 10, or 20 years in the controls in the Sanderson et al[1] study (Table 3). Note that log transformation of the lagged exposure variable markedly influences the product moment correlation.

Case Example

The Sanderson et al[1] study of beryllium exposure and lung cancer was a case–control study nested within an occupational cohort. The cohort consisted of subjects who had worked at least 2 days in a beryllium materials production factory between January 1, 1940, and December 31, 1969. This cohort was followed through 1992 for mortality with 142 deaths from lung cancer being identified. For each lung cancer death, five controls were randomly selected from the set of study subjects who had a *date of hire* less than the *age at censor* (death) of each case and an *age at censor* equal to or greater than the case's *age at censor*. The controls for each case had their experience truncated at the case's *age at censor*. Exposure metrics (tenure, cumulative, average, and maximum) were created for lag times of 0, 10, and 20 years. The analysis was conducted in two ways. One was to compare the geometric means of the exposure variables in cases and controls. The other was to use conditional linear logistic regression to identify odds ratios associated with exposure quartile categories. This design could be represented by a causal diagram[8] (Fig. 5). First, unmeasured factors (X) create an association between

TABLE 4. This Table Shows the Odds Ratios From a Conditional Logistic Regression Analysis of Quartiles of Cumulative Beryllium Exposure Lagged 10 Years With the Effects of Adding to the Analysis Timescales. It Also Shows the Product Moment Correlation of the Time Scales With *Date Of Birth* and *Time From Hire* in the Controls in the Study[1]

Time Scale	Odds Ratios for Quartiles of Cumulative Exposure Lagged 10 Yrs				Correlation With	Correlation With	Odds Ratios for Quartile of Control Variable			
Control Variable	1	2	3	4	*Date of Birth*	*Time From Hire*	1	2	3	4
None	1	1.34	1.39	0.95						
Date of birth	1	1.23	1.27	0.85		0.59	1	2.07*	2.20*	2.08*
Time from hire	1	1.16	1.19	0.79	0.59		1	1.77	2.48*	2.45*
Age at hire	1	1.20	1.19	0.80	−0.91	−0.73	1	0.79	0.62	0.36*
Date of hire	1	1.34	1.39	0.94	0.44	−0.16	1	0.92	1.01	1.00
Date of censor	1	1.14	1.20	0.81	0.77	0.89	1	2.60*	2.99*	2.40*

*Confidence limits do not include 1.0.

age at censor and lung cancer. *Age at censor* was considered a potential confounder and was controlled via the *age at censor* risk set matching procedure for controls. Second, exposure profiles were modified using *lag time* to create several lagged exposure variables as putative causes of lung cancer. Schubauer-Berigan et al[3] added a second time scale to the causal diagram (Fig. 5) on the basis that unspecified factors (Y) created a demonstrable association between the *date of birth* and lung cancer. *Date of birth* was considered to be a confounder of the lagged exposure—lung cancer relationship and addition of *date of birth* to the analysis, using quartile categories, importantly changed the odds ratios for lagged exposure categories. For instance, for cumulative exposure lagged 20 years, the estimates for odds ratios for quartiles of cumulative exposure were 1.0, 2.18, 1.89, and 1.89 from lowest exposure to highest without control of *date of birth* and were 1.0, 1.46, 1.29, and 1.30 after adjustment with *date of birth* quartiles. *Age at hire* quartile categories alone (without *date of birth* in the model) changed the cumulative exposure categories, lagged 20 years, slightly more, to 1.0, 1.37, 1.21, and 1.19, respectively. No separate rationale was developed for *age at hire* being a confounder other than the correlation between *date of birth* and *age at hire* of −0.906. No open association path was specified for either *date of birth* or *age at hire* that linked to both lung cancer and lagged exposure.

We added two new features to this causal diagram. First, we added *time from hire* as a third determinant of lagged exposure (Fig. 5). This variable was added as the result of the demonstration that lagging exposure creates an association between lagged exposure and *time from hire* (Table 3). Second, we looked at the time scale web (Fig. 2) to identify potential time scale association paths linking *date of birth* with *time from hire*.

Date of birth is in three relational triads, so there are six time scales, which could form a one-stop path from *date of birth* to *time from hire.* Two, *age at termination* and *date of termination*, do not share a relational triad with *time from hire*. Among the four that do share a relational triad with *time from hire*, *age at censor* is already controlled via control selection and truncation. In the Sanderson et al[1] study, control subject's *date of hire* is weakly correlated with both *date of birth* and *time from hire*, whereas both *age at hire* and *date of censor* are strongly correlated with both (Table 4). *Age at hire* is also correlated with *date of censor* (−0.72). Adjustment of the odds ratios for quartiles of cumulative exposure lagged 10 years with quartiles of *date of birth* reduces the odds ratios. Adjustment with quartiles of *time from hire*, *age at hire*, and *date of censor* reduces the cumulative exposure odds ratios further. Consistent with its low correlation with *date of birth* and *time from hire*, adjustment with *date of hire* has no effect on the cumulative exposure lagged 10 years odds ratios. These findings support the hypotheses that two paths in the time scale web, lung cancer—*date of birth—age at hire—time from hire*—lagged exposure and lung cancer—*date of birth—date of censor—time from hire*—lagged exposure, function as important association paths (Fig. 5) contributing to confounding.

DISCUSSION

The lessons from the structure of occupational cohort time scales and the case example are several. First is that, in addition to developing a rationale for possible confounders through their association with disease, it is also worthwhile to look for possible confounders through association with exposure. Lagging exposure modifies the association of time scales *time from hire* and *time from termination* with exposure. Log transformation of exposure variables can modify these associations.

Second is that a time scale web is created by the relational triads and their interconnections (Fig. 2). The many paths connecting time scales in the web are always present in any occupational cohort study. The question in any study is "Can an open path connect exposure with disease, and if so, how long is the path and how large are the associations in each step of the path?" We found it useful to calculate a correlation matrix between all 10 time scales and the exposure variables. The usefulness of this increased greatly when we could place the correlations in the context of relational triads and their connections in the web.

Third is that both the relational triads and their web structure imply that with more than one time scale in an analysis or otherwise controlled, other time scales may be involved in the analysis. Specific values for any two time scales in a relational triad will calculate a value for the third. Thus, if two are in an analysis, for instance, *date of birth* and *age at censor*, so is the third, *date of censor*. Extending this, if certain combinations of three time scales are in an analysis, two in one relational triad and one in a connected relational triad, a total of six time scale variables can be calculated and are therefore represented. For example, with *date of birth* and *age at censor*, which imply a value for *date of censor*, the addition of *date of hire* implies values for *age at hire* and *time from hire*, for a total of six time variables included. Addition of any one of the remaining four time scale variables allows the calculation of all 10, implying that with certain combinations of four time scales in an analysis all 10 are included.

It can be inferred from the property that certain combinations of four time scale variables will calculate values for all 10 that there are only four pieces of information among the 10. This conclusion seems obvious when the four time scales are, for example, *date of birth*, *date of hire*, *date of termination*, and *date of censor*, or *date of birth*, *age at hire*, *age at termination*, and *age at censor*, but seemingly odd combinations also suffice, for instance, *tenure*, *time from termination*, *date of censor*, and *age at termination*, The difference between the first two and the third is that in the first two all 10 time scales can be calculated directly from the four, whereas in the third intermediate calculations are required.

Besides obscure confounding, there is another mechanism through which relational triads and their connections may complicate epidemiologic analyses. In a second example, the approach to analysis of occupational cohort data was to stratify by age and date in 5-year intervals and assess risk per person year within these restricted categories.[9] As one definition of *age at censor* and *date of censor* is the age and date used in an analysis, we will use these terms. Constraint of *age at censor* and *date of censor* to 5-year intervals constrains possible *date of birth* to 10-year intervals. Using Fig. 2 to identify the six relational triads connected to *age at censor*, *date of censor*, and *date of birth*, three contain combinations of *date of hire*, *age at hire*, and *time from hire*, and three contain *combinations of date of termination, age at termination*, and *time from termination*. Taking the former, it would be expected that *date of hire*, *age at hire*, and *time from hire* would be very highly correlated with each other. Suppose then one performed analyses of disease risk with each and found a sharp rise in disease risk with increasing *time from hire*, a sharp fall in risk with increasing *date of hire*, and a sharp fall in risk with increasing *age at hire*. One could interpret the first as a latency effect, suggesting exposure-disease causality, the second as (assuming falling levels of exposure with time) an exposure-response causal effect, and the third as an age-susceptibility effect, also supporting an inference of a causal effect of the workplace exposure. The high correlation between the three time scale variables *time from hire*, *date of hire*, and *age at hire* ensures that the three analyses are not independent of one another and renders ambiguous any inference about which association was the driver of the effect.

It is not generally recognized how lagging exposure and log transformation of exposure variables can modify associations of *time from hire* and *time from termination* with exposure variables. Exploratory analysis with different lag times and with untransformed and log transformed exposure variables, where demonstration of an association between exposure and disease is the desired outcome, carries the risk that either or both combined can activate association paths in the time scale web and introduce obscure confounding. This is the lesson of the case example. In an analysis,[1] in which geometric means of exposure variables were compared between cases and controls, the combination of lagging and the log transformation implicit in geometric means activated the *time from hire − age at hire/date of censor − date of birth* confounding paths also. Very different results and interpretations would have been reached if means of un-log transformed exposure variables had been compared (data not shown). Similarly, lagging alone creates strong rank correlation of *time from hire* with lagged exposure and thereby activated the *time from hire − age at hire/date of censor − date of birth* confounding paths for ordinal category analyses. Because these paths were not obvious, when the desired outcome, demonstration of association between exposure and disease was observed it was accepted.

The mathematics of relational triads are precise when time scales are continuous variables. Time scales are often collapsed into ordinal categories, which would affect the precision of the relationships we have examined. We have not explored the behavior of relational triads when the time scales are categorical. For example, when *date of birth* and *age at hire* are in an analysis as categorical variables, to what extent is *date of hire* represented? Are residual effects small and unimportant or are there poorly understood but important consequences?

We also have not explored how the effects of constraint within strata of a variable translate to the overall effects. For example, in Sanderson et al,[1] within case–control sets the correlation between *time from hire* and *age at hire* is −1.0, whereas across all the controls it is −0.73. We are unclear which value, −1.0 or −0.73, best characterizes the strength of the association for the purposes of understanding the strength of the association path.

Use of the Occupational Cohort Time Scale Web to Plan Analyses

The property that inclusion of certain combinations of time scales implies the inclusion of others can be used to plan analyses to maximize efficiency and avoid redundancy. Identification of hidden time scales can be accomplished by circling in the time scale web diagram the time scales desired to be included, determining which form a side on a relational triad, and identifying time scales, which occupy the third corner, and circling these, and so on. An example where this seems to have been done is an occupational cohort study of diesel exhaust carbon particle exposure and lung cancer,[10] which clearly described the reason for including each time scale variable. In a proportional hazards model, *date of hire*, *date of censor*, *age at censor*, and *tenure* were included as adjustment variables. From the time scale web it can be deduced that these four allow the calculation of three more, and with these, all ten. Furthermore, with *tenure* in the model, when either *cumulative exposure* or *average exposure* was examined, the multiplicative exposure relational triad included the other. Thus, five variables visible in the model include all 12 in Fig. 2, and there is no redundancy.

The occupational cohort time scale relational triads and web can be used to identify potential cofounding. First, prepare a list of time scales likely associated with disease or exposure including when lagging exposure *time from hire* and *time from termination*. Then, circling these on Fig. 2 and identify connecting paths. These paths can be evaluated by calculating all 10 time scales for each subject and then constructing a correlation matrix containing all time scale variables and all exposure variables, including lagged variables. These correlations estimate the strength of each link in a path.

The next layer of inquiry is within the unit of analysis, for example, in the case example a case and its five controls, and in the second example sets of person years within sets bounded by 5-year by 5-year categories of age and date. The constraints inherent in the construction of the unit of analysis can be applied to the relational triads identified in the time scale web to infer partial correlations within the unit of analysis that may be stronger or weaker than those in the overall data. These inferred correlations can be used to evaluate potential association paths or ambiguity of interpretation.

Use the time scale web to evaluate how published papers handled potential time scale confounding or ambiguity of interpretation

For the occupational physician wishing to critically evaluate a published paper with respect to time scales, a number of questions can be asked. The answers to these questions can give insight into the strengths and weaknesses of the paper.

- Does the methods section describe which time scales are included in the analyses and why? If this is not clear, it is quite possible the authors were unclear in their approach.
- If exposure was lagged, does the paper indicate whether *time from hire* and *time from termination* were considered as potential confounders?
- What is the unit of analysis and which time scales define that unit? Given these visible time scales, what hidden time scales can be deduced from the time scale web (Fig. 2) to be included in the analysis?
- What are the constraints on the visible and hidden time scales in the analysis and what correlations can be inferred for the time scales forming the rest of the relational triads?
  - What time scales are possible sources of ambiguity of interpretation?
  - What time scales not in the analysis (neither visible nor hidden) remain as potential confounders?

CONCLUSIONS

Occupational cohort time scales can be arranged as interconnected relational triads. Understanding the properties of these relational triads and of the connecting time scale web is helpful in understanding the origins of otherwise obscure, but possibly strong, confounding bias, as well as ambiguities of interpretation. Consideration of the interrelations between the occupational cohort time scales contributes to rational planning of use of time scales in analyses. Recognition of the effect of lagging exposure and log transformation on the association of time scales *time from hire* and *time from terminations* with exposure aids identification of open association paths.

REFERENCES

1. Sanderson WT, Ward EM, Steenland K, Peterson MR. Lung cancer case–control study of beryllium workers. *Am J Indust Med*. 2001;39:133–144.
2. Deubner DC, Roth HD, Levy PS. Empirical evaluation of complex study designs: occupational exposure and cancer. *J Occup Environ Med*. 2007; 49:953—959.
3. Schubauer-Berigan MK, Deddens JA, Steenland K, Sanderson WT, Petersen MR. Adjustment for temporal confounders in a reanalysis of a case–control study of beryllium and lung cancer. *Occup Environ Med*. 2008;65:379–383.
4. Rothman KJ, Mosquin PL. Confounding after risk set sampling in the beryllium study of Sanderson et al. *Ann Epidemiol*. 2011;21:773–779.
5. Ward E, Okun A, Ruder A, Fingerhut M, Steenland K. A mortality study of workers at 7 beryllium processing plants. *Am J Indust Med*. 1992;22:885–904.
6. Beaumont JJ, Steenland K, Minton A, Meyer S. A computer program for incidence density sampling of controls in case–control studies nested within occupational cohort studies. *Am J Epidemiol*. 1989;129:212–219.
7. Rothman KJ: Induction and latent periods. *Am J Epidemiol*. 1981;114:253–259.
8. Glymour MM, Greenlkand S. Causal diagrams. In: Rothman KJ, Greenland S, Lash LL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Williams; 2008:183–212.
9. Schubauer-Berigan MK, Couch JR, Petersen MR, Carrion T, Jin Y, Deddens JA. Cohort mortality study of workers at seven beryllium processing plants: update and association with cumulative and maximum exposure. *Occup Environ Med*. 2011;68:345–353.
10. Garshick E, Laden F, Hart JE, Davis ME, Eisen EA, Smith TJ. Lung cancer and elemental carbon exposure in trucking industry workers. *Environ Health Perspect*. 2012;120:1301–1306.

