COMMENTARY

Shift work and cancer research:
A thought experiment into a potential chronobiological fallacy of past and perspectives for future epidemiological studies

Thomas C. Erren 1, Peter Morfeld 2

1 Institute and Policlinic for Occupational Medicine, Environmental Medicine and Prevention Research, UNIKLINIK KÖLN, University of Cologne, Germany.
2 Institute for Occupational Epidemiology and Risk Assessment (IERA), Evonik Industries AG, Germany.

Correspondence to: Prof. Thomas C. Erren, MD., MPH.
Director of the Institute and Policlinic for Occupational Medicine, Environmental Medicine and Prevention Research, University of Cologne, Lindenthal, Germany.
TEL: +49-221-4784450; FAX: +49-221-4785119; E-MAIL: tim.erren@uni-koeln.de

Submitted: 2013-05-31 Accepted: 2013-06-06 Published online: 2013-06-25

Key words: chronobiological fallacy; thought experiment; chronotype; chronodisruption; internal time; external time; shift work; circadian disruption; cancer; health; disease; prevention

Abstract
With their 2007 classification – shift work involving “circadian disruption” is probably carcinogenic to humans (Group 2A) – the International Agency for Research on Cancer [IARC] provided a riddle for scientists and the public alike. Thereafter, eighteen epidemiological investigations into shift work and a host of malignant endpoints (including cancers of the breast, prostate, lung, colon, rectum, pancreas, bladder, skin and non-Hodgkin lymphoma [NHL]) as well as mortality were published. Although IARC experts identified “circadian disruption” as the critical link in the “probable” chains of cancer causation, almost none of the post-IARC studies specifically considered a disturbed temporal organization of biology. This implies that epidemiological research to-date is less focused than it should be. To illustrate a potential chronobiological fallacy of past studies, we offer a thought experiment. In addition, we consider first empirical evidence from recent research which avoided such bias. Methodological perspectives for future chronobiology-driven epidemiological research are outlined.

INTRODUCTION

In 2007, the International Agency for Research on Cancer [IARC] classified shift work involving “circadian disruption” [CD] as a probable human carcinogen (Group 2A) (Straif et al. 2007; IARC 2010). Two lines of evidence formed the basis for this conclusion: there is “sufficient evidence in experimental animals for the carcinogenicity of light during the daily dark period (biological night)” and “limited evidence in humans for the carcinogenicity of shift work that involves night work” (Straif et al. 2007).

THE PROBLEM

While the observational studies may overall be interpreted as being in line with the “probable” link between shift work and cancer, none of them assessed the critical link in the postulated chain of causation, i.e., “circadian disruption” and almost none specifically considered critical facets of the temporal organization of biology such as chronotype or chronobiological propensity, internal time and chronodisruption (Table 1).

And yet, at and beyond IARC, experts agree that disturbed temporal organization is at the core of concern that shift work, operationalized as work at chronobiologically unusual times, may put women and men at risk of developing cancer. In the following, we present a fictitious scenario to explore the likely consequences of what we consider a potential chronobiological fallacy of prevailing epidemiological research. More specifically, this Gedankenexperiment contributes to answering the following questions:

(i) Why could hitherto conducted studies regarding shift work and cancer be affected by a potential chronobiological bias and fallacy?
(ii) Is there empirical evidence from shift work and cancer research which avoided the potential chronobiological fallacy by considering chronobiological propensity in assessments of cancer risk?
(iii) How could this potential chronobiological fallacy be avoided in much-needed future studies?

CHRONOBILOGICAL FALLACY

In our view, studies of shift workers and cancer risks which do not compare an individual’s “internal time” with the “external times” forced upon him or her by shift work or other activities lead to a logical fallacy. Some principal considerations may clarify why we suggest to work with the term “chronobiological fallacy”.

When we expect specific stress and strain in workers who are engaged in shift work compared to those who are not, it is the time window when such work is required and done which we consider as causative. More generally, we tend to expect that work during the day is less demanding than at night. Moreover, we expect that individuals who rotate with their shifts through day and night experience disturbed biological rhythms. Abundant research has unambiguously demonstrated that workers possess an individual propensity to be awake and asleep in 24 hour time windows (Roenneberg et al. 2003; 2007). With this background, we must compare the genetically determined chronotype as indicator of the “internal time” with the “external time” of the shift work associated time windows to assess disturbed biological rhythms. In this vein, it was proposed to use the individual chronotype as a tool to approximate a person’s individual susceptibility to working at biologically unusual times which cause most stress and strain at the individual level (Erren 2010; 2013). Most studies to-date failed to specifically consider exposure to, and possible doses and gradients of, CD. To avoid this, future studies must compare a given individual’s chronotype with

<table>
<thead>
<tr>
<th>Publication</th>
<th>circadian disruption</th>
<th>chronotype</th>
<th>internal time</th>
<th>chronodisruption</th>
</tr>
</thead>
<tbody>
<tr>
<td>1972–2007 Pre-IARC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor (1972)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tynes (1996)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davis (2001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hansen (2001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schernhammer (2001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schernhammer (2003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lie (2006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O’Leary (2006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schernhammer (2006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kubo (2006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwartzbaum (2007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viswanathan (2007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conlon (2007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007–2013 Post-IARC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lahti (2008)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pukkala (2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oberlinner (2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pesch (2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pronk (2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kubo (2011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schernhammer (2011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lie (2011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nätti (2012)</td>
<td></td>
<td>yes</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Hansen (2012a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hansen (2012b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent (2012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yong (2013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knutsson (2013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Menegaux (2013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rabstein (2013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatti (2013)</td>
<td></td>
<td>yes</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Lin (2013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the timing of his/her shifts to answer the key question “How much is the physiological nexus between internal and external times disrupted?” (chronodisruption: Erren & Reiter 2013). Comparing chronotype information which we obtain for instance via the Munich ChronoType Questionnaire for Shift-Workers (MCTQshift; Juda et al. 2013) with the shift work associated times promises important information regarding an individual’s very susceptibility and exposure to, and dose of, “circadian disruption”.

Disconcertingly, the potential chronobiological fallacy may have masked both the detection and very magnitude of possible cancer risks associated with work in time windows which are chronobiologically unusual for a given individual.

That “circadian disruption” was not specifically assessed in targeted studies after 2007 may be somewhat explained by IARC itself. In fact, despite its identification as critical link in the “probable” chain of cancer causation, “circadian disruption” is mentioned merely 12 times on 203 pages in the IARC monograph 98 on “shift work” and neither a definition nor a way to approximate CD in observational research was provided (IARC 2010).

THOUGHT EXPERIMENT

Regarding our question (i), consider scientists sitting on the edge of a hill overlooking a large Factory A. Their hypothesis is that workers entering A are exposed to doses of some carcinogen. If the scientists were to talk with the workers they would find out that individuals in A are – at times – provided with protective gear to shield them from exposures which are – at times – carcinogenic but otherwise not. Without information on these two time windows, a slight increase of cancer risk detected in those who work in A would have to be indiscriminately attributed to all workers in the “black box” (Savitz 1994) of A.

But rather than missing the exact nature and magnitude of such signals from the “black-box” of A, the scientists could ask individual workers key questions: First, when are you provided with protective gear in A (T1)? Second, when are you exposed to the carcinogen in A (T2)? These two times, T1 and T2, critically determine the risk of cancer observed in workforce A. As long as T1 and T2 are in phase (T1 = T2), workers are not at increased risk to develop cancer. However, as long as we lack information on both time windows T1 and T2, all workers leaving A must be considered erroneously as exposed there, at all times. But only those workers for whom the nexus of T1 and T2 splits (T1 ≠ T2) are at an increased risk of developing cancer and this risk is likely to increase the more the longer the time window dissociate or the nexus of the two times disrupts (Table 2; Erren & Reiter 2013). Without comparing T1 and T2 at the individual worker level, the results when comparing A’s workforce with a workforce in Factory B with no carcinogenic exposure must be expected to be dire and misleading: In fact,

- An increased cancer risk may be detected in workforce A;
- Cancer risks in A would be overestimated for those who were not exposed to relevant carcinogenic doses;
- Cancer risks in A would be underestimated for those who were exposed to relevant carcinogenic doses.

Put differently: Temporal organization of biology is at the heart when weighing evidence for or against links between shift work involving “circadian disruption” and cancer. When two times are critically involved, namely “internal time” determined by the chronotype and “external time” determined by the shift work regimen, we need to compare the two rather than relying on one of them (Table 2).

EMPIRICAL EVIDENCE

Regarding question (ii), two recent studies may have opened doors to the next level of epidemiological research into shift work involving circadian disruption and biologically plausible cancer risks (IARC 2010).

Different from all other studies in this important field of research, only Hansen and Lassen (2012) and Bhatti et al. (2013) explicitly considered chronobiological propensity as a means to zero in on a chronotype-associated susceptibility to night-shift and rotating-shift work (Erren 2013). In the study on female night-shift workers in the Danish military (Hansen & Lassen 2012), breast cancer risks in morning types and evening types were almost quadrupled and doubled, respectively, while they were not increased in the “neither” diurnal preference category. When Bhatti et al. (2013) studied female night-shift workers, they found evidence that ovarian cancer was elevated to a higher extent in morning types than in evening types and suggested that future studies should include detailed assessments of diurnal preference, i.e., chronotype.

Tab. 2. Effects of temporal relationships between two critical time windows: Thought experiment and shift work scenarios.

<table>
<thead>
<tr>
<th></th>
<th>Thought Experiment</th>
<th>Shift work scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>protective gear</td>
<td>biological day</td>
</tr>
<tr>
<td>T2</td>
<td>carcinogen exposure</td>
<td>shift work associated times</td>
</tr>
<tr>
<td>T1 = T2</td>
<td>cancer risks: no</td>
<td>CD*: no</td>
</tr>
<tr>
<td>T1 ≠ T2</td>
<td>cancer risks: yes</td>
<td>CD*: yes</td>
</tr>
</tbody>
</table>

*CD: circadian disruption (undefined; IARC 2010); chronodisruption (defined; Erren et al. 2003; 2008; 2009; 2013)
METHODOLOGICAL SOLUTION

Regarding question (iii), how could we tackle the suggested bias and fallacy which is generated by failing to consider key chronobiological insights in epidemiological settings?

The described problem is caused by applying an inappropriate exposure metric. This problem is well known to occupational epidemiologists and sometimes described and attacked by distinguishing between “exposure” and “dose” (Checkoway et al. 2004). Whereas the dose is per definition assumed to be proportional to risk, exposure may fail to have such a simple relationship with the outcome of interest. Using exposure instead of dose can generate severe misclassification. This can lead to underestimated or missed associations (Seixas & Checkoway 1995).

In general, the relevant dose metric can be derived by weighing exposure periods within each subject based on formal hypotheses about underlying pathophysiological processes (Kriebel et al. 2007). Richardson et al. (2011) investigated how to estimate appropriate weighting schemes in latency analyses. They discussed that fixed schemes without individual variation may be inappropriate. However, without having additional external information to address this point, their approach was restricted to applying random lagging periods. In our case, we may use $T_1$ and $T_2$ as additional information to derive an individual time-dependent weighting scheme $W$. $T_1$ and $T_2$ in real life correspond to an individual’s biological day and to his/her time window of shift work or other activities, respectively (Table 2). We can obtain information on these internal and external times for instance via the novel Munich ChronoType Questionnaire for Shift-Workers (MCTQshift; Juda et al. 2013) and via questions regarding an individual’s shift work schedules and further associated activities such as preparing for and getting to and from work. The weight $W$ can then be applied to modify the exposure $T_2$ into a more appropriate dose metric $D$. The individual $W$ is defined as 1 but set to zero when $T_1$ and $T_2$ are in phase. The individual time-dependent $D$ is then derived by multiplying $W$ and $T_2$ at each point in time, i.e., only those parts of $T_2$ are kept within the dose estimate that are not in phase with $T_1$. This procedure defines a process $D$ within each subject under study. The definition of $D$ can be refined by not using a binary $W$ but a score $S$ as weight, $0 \leq S \leq 1$, that reflects how far $T_1$ and $T_2$ are out of phase at a specific point in time within this individual. $D$ may be integrated across time within each subject to derive an individual time-dependent cumulative dose measure, estimating the cumulative amount of “circadian disruption” the individual experienced due to $T_1$ and $T_2$. Finally, this dose measure can be used in epidemiological models to study the relationship with cancer incidences or mortalities while adjusting for covariates (Rothman et al. 2008).

CONCLUSION

It is imperative to avoid the illustrated potential chronobiological fallacy in much-needed targeted observational studies to ‘test’ the “probable” causal links between shift work, disturbed biological rhythms and cancer. More details on how we may consider the split nexus of internal and external times in studies of shift workers and whatever endpoint have been provided elsewhere (Erren & Reiter 2013).

Taken together, the inclusion of chronobiological propensity or chronotype as temporal markers (chronomarkers: Erren & Reiter 2013) of susceptibility to, or dose of, “circadian disruption” or chronodisruption by Hansen & Lassen (2012) and Bhatti et al. (2013) is commendable. These studies should serve as a prelude to research which must evolve to achieve two ends: To rigorously meet the 2007 IARC challenge, i.e., to exonerate or understand chronobiologically-plausible links between shift work – including night work – and cancer, on one hand, and to identify means to break possible chains of causation, on the other.

It does not escape our attention that avoiding the potential chronobiological fallacy can not be confined to studying possible effects of shift work on the development of internal cancer. It is indeed a conditio sine qua non for any observational study which investigates possible links between shift work and effects on health and disease in general.

Conflict of Interest

None.

REFERENCES


