Accident Analysis and Prevention 99 (2017) 469–473

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

How should a bio-mathematical model be used within a fatigue risk management system to determine whether or not a working time arrangement is safe?

Drew Dawson*, David Darwent, Gregory D. Roach

Appleton Institute for Behavioural Science, Central Queensland University, PO Box 42, Goodwood, SA 5034, Australia

ARTICLE INFO

Article history:
Received 9 September 2015
Received in revised form 29 October 2015
Accepted 26 November 2015
Available online 31 March 2016

Keywords:
Sleep
Wake
Circadian rhythms
Sleepiness
Fatigue model
Shiftwork

ABSTRACT

Bio-mathematical models that predict fatigue and/or sleepiness have proved a useful adjunct in the management of what has been typically referred to as fatigue-related risk. Codifying what constitutes appropriate use of these models will be increasingly important over the next decade. Current guidelines for determining a safe working time arrangement based on model outputs generally use a single upper threshold and are, arguably, over-simplistic. These guidelines fail to incorporate explicitly essential aspects of the risk assessment process – namely, the inherent uncertainty and variability in human sleep–wake behavior; the non-linear relationship between fatigue, task performance and safety outcomes; the consequence of a fatigue-related error and its influence on overall risk; and the impact of risk mitigation or controls in reducing the likelihood or consequence of a fatigue-related error. As industry and regulatory bodies increasingly move toward performance-based approaches to safety management, any fatigue risk management system that includes a bio-mathematical model should specify what exactly is measured by the model, and how the model can be used in the context of a safety management system approach. This will require significant dialog between the various parties with an interest in bio-mathematical models, i.e. developers, vendors, end-users, and regulators.

1. Introduction

In recent years, risk-based approaches to the management of fatigue have evolved as a viable, if not desirable, alternative to compliance-based approaches (Dawson and McCulloch, 2005; Gander et al., 2011). Organizations and regulators have both advocated this approach because it potentially provides a more sophisticated method for better identifying safe (or unsafe) working time arrangements (WTA), potentially improving safety and increasing operational flexibility and productivity.

In risk-based fatigue management programs, the WTA for a group of workers is typically risk-assessed using a standardized methodology (e.g. AS 4360/ISO 31000). In general terms, this methodology quantifies the risk using the arithmetical product of the *likelihood* and *consequence* of a fatigue-related error. The assessment then evaluates whether the mitigations/controls in place are sufficient to reduce the risk associated with a given WTA to a level considered acceptable to the community (or the regulator as proxy thereof).

A key element of this shift in regulatory approach has been the introduction of bio-mathematical models (BMMs) of fatigue (e.g. Mallis et al., 2004). These models purport to predict the general construct of fatigue based on sleep–wake behavior and/or the WTA. It is worth noting that the term fatigue has been used somewhat loosely in this discourse and is often substituted for the more biologically precise term of sleepiness or the more technically correct term of sleep opportunity. Developers and vendors of models have typically used the term 'fatigue' to describe the models in response to regulatory requirements and the common-language use of the term in industry. In this paper, we have typically used the terms 'fatigue' and 'sleepiness' somewhat interchangeably.

Compared with more traditional approaches in which safety is inferred from compliance to a prescriptive rule set of shift and break maxima and minima, BMMs have been advocated as a more reliable and valid way to determine the level of risk associated with a WTA and, to a certain extent, whether it can be considered safe or not (Dawson and McCulloch, 2005). BMMs have been used primarily to quantify the degree of sleep opportunity afforded by a roster or schedule and, by inference, the *relative likelihood* of a

* Corresponding author.
E-mail addresses: drew.dawson@cqu.edu.au (D. Dawson), david.darwent@cqu.edu.au (D. Darwent), greg.roach@cqu.edu.au (G.D. Roach).

http://dx.doi.org/10.1016/j.aap.2015.11.032
0001-4575/

fatigue-related error (Dawson et al., 2011). In conjunction with an assessment of the consequence of a fatigue-related error and the mitigations in place, the net risk can then be determined in a semi-quantitative manner.

2. Determining a safe working time arrangement using bio-mathematical models

There are several BMMs used for predicting fatigue, sleepiness, sleep opportunity, fatigue likelihood, etc. that are either commercially available [e.g. FAID (Roach et al., 2004), SAFE (Belyavin and Spencer, 2004), SAFTE (Hursh et al., 2004)] or publically available [Three Process Model of Alertness (Ingre et al., 2014)]. While all of these models have been validated to some extent, only the Three Process Model of Alertness is currently in a position to be independently validated because both its algorithms and parameterizations have been published. A common feature of these different software tools has been the use of a threshold value to designate a 'safe level' of operations with respect to fatigue-related risk. Some of the models have this as an inbuilt 'feature' of the software; others require it as an input variable based on a formal risk assessment process. In other cases, external parties such as scientists, regulators or consultants have determined these thresholds based on a combination of evidence, experience and political expediency. While intuitively appealing, the evidence to support many of the advocated thresholds is limited and, with few exceptions, neither strongly evidence-based nor consistent with a risk/safety systems approach.

As the use of BMMs increases, the question of how to determine a 'safe' working time arrangement will become increasingly important. The modal approach at the moment, which is to specify a threshold, is really a vestigial remnant of the 'culture of prescription' around WTAs – merely substituting threshold values for shift maxima and break minima with a comparable threshold value for a BMM. While understandable from a naïve perspective, the process whereby organizations and regulators determine acceptable WTAs using BMMs will be subject to increasing scrutiny – especially in the context of accidents and subsequent litigation.

The currently favored approach, i.e. compliance above or below a simple fatigue likelihood threshold, is unlikely to be considered legally or scientifically defensible. Moreover, as regulatory models move increasingly toward a risk- and safety-management systems approach (Gander et al., 2011), a threshold-based approach to the use of BMMs is unlikely to be considered consistent with the requirements of a risk-based approach.

Compared to compliance with prescriptive hours, there is little doubt that BMMs significantly improve our capacity to predict the likelihood of fatigue across any given WTA. This is because model algorithms are optimized against observed fatigue data and inputs include better predictor variables, such as [estimated] prior sleep–wake history, time-of-day, etc. (Friedl et al., 2004). However, better prediction of fatigue [likelihood] does not automatically ensure a safer workplace. If we are to realize the full potential of BMMs as an integral component of fatigue risk management systems or safety management systems, we will also require a more sophisticated, risk-based approach to the judgment of what is, or is not, a safe WTA.

3. How do we define what is safe using a bio-mathematical model?

In determining what constitutes a safe WTA, it is important to first 'unpack'

(a) what a BMM actually measures,
(b) the predictive relationship between a BMM output and task safety/performance, and
(c) the most appropriate way to determine what is considered 'nominally safe' using a BMM.

3.1. What does a bio-mathematical model measure?

It has previously been argued that BMMs fall into two classes – one step and two-step models (Kandelaars et al., 2006). One-step models use prior sleep–wake history to predict fatigue or sleepiness. These models have a long history (e.g. Daan et al., 1984) and with the exception of some relatively artificial sleep–wake schedules, are reasonably accurate and reliable predictive tools (McCauley et al., 2013).

Unfortunately, most organizations that use BMMs do not have access to an individual worker's sleep–wake history. Model developers and vendors have 'finessed' this problem using a two-step BMM. Two-step models use the timing and duration of shifts to estimate sleep–wake history for an 'average employee'. The estimated 'average' sleep–wake history then forms the input for subsequent 'fatigue' or 'sleepiness' prediction for a work group using a standard one-step model (*as per above*). Two-step models homogenize a work groups' sleep–wake histories by eliminating inter- and intra-individual differences and, as a consequence, create an artificial consistency to sleep–wake behavior.

While homogenizing sleep–wake behavior is problematic at the individual or event level, at the aggregate level it enables a global risk assessment to be conducted before a roster is worked. The predictions will, however, lack specificity and sensitivity at the individual worker or specific event level. In general, it is probably more accurate to suggest that a two-step BMM estimates fatigue likelihood based on average sleep–wake behavior derived from the sleep opportunity associated with a WTA. These BMM can be augmented by cross-referencing outputs with pre-existing distributions of sleep (i.e. mean sleep and percentile distributions of sleep), showing how much sleep work groups obtain when working shifts in the same output range (Darwent et al., 2015). The number of employees obtaining more or less sleep than the mean provides an indication of the number of employees who would likely be more or less alert than predicted, respectively.

3.2. The link between bio-mathematical model outputs and task safety/performance

It is worth noting that fatigue is not yet a directly observable or measurable phenomenon, but rather, it is typically inferred from indirect measures (Baulk et al., 2008; Dinges et al., 1998). These measures include (a) signs and symptoms of fatigue based on individual physiology, e.g. yawning, changes in ocular behavior, changes in EEG, etc. (Dawson et al., 2014), (b) impairment of cognitive performance derived from computer administered tasks specifically designed to be sensitive to the effects of the known determinants of fatigue, i.e. prior sleep–wake history, time-of-day, and time-on-task (Basner and Dinges, 2011), and (c) self-report measures of sleepiness and fatigue such as the Karolinska Sleepiness Scale (Åkerstedt and Gillberg, 1990), the Stanford Sleepiness Scale (Hoddes et al., 1973), and the Samn–Pereli Fatigue Scale (Samn and Perelli, 1982).

While embedded performance measures, such as steering lane variability, gear changes, braking behavior, and accelerating behavior, provide a more ecologically valid measure of task performance, it is especially important to note that 'performance' on physiological and/or laboratory-based tasks may not be a reliable proxy for [safe] task performance in the workplace (Dawson et al., 2014). Despite the self-evident nature of this caveat, some of the commercially available BMMs provide graphical outputs with 'performance'