FATIGUE: IDENTIFICATION, MANAGEMENT AND COUNTERMEASURES

Mark S. Young and Tabitha Steel

Rail Accident Investigation Branch, UK

Managing operator fatigue is a perennial challenge in any safety-critical industry that involves shiftwork. Whilst the causes and effects of fatigue are well documented, there have been a number of recent rail accidents in which monitoring, reporting and mitigating the risk arising from fatigue have played a part. This paper presents a review of literature and practices in rail and other safety-critical industries, aimed at understanding barriers to reporting as fit for duty when possibly suffering from fatigue, as well as identifying and managing the accumulation of fatigue over the duration of a shift.

Introduction

The demands of a 24/7 society mean that many jobs now involve an element of shiftwork. With shiftwork comes the problem of fatigue, since humans are diurnal animals and are not adapted to working through the night.

Fatigue is a factor of three processes in combination (e.g., Ingre et al., 2014): circadian rhythm, the amount of time spent asleep, and the time since waking up. In shiftwork, some of the associated risk factors are length of shift, rest between duties (Härmä et al., 2002), shift rotation and disrupted sleep (Åkerstedt, 1998).

The short-term effects of fatigue include not just increased sleepiness, but also impaired decision-making, degraded task performance and, ultimately, an increased risk of errors and accidents (RSSB, 2012).

The UK rail industry is not immune to the problem of fatigue and, although considerable efforts have been made to mitigate the risk of fatigue in recent years, it still plays a part in a number of rail accidents. The Rail Accident Investigation

Branch (RAIB) has investigated 12 accidents that included fatigue as a factor since it became operational in October 2005. A key issue emerging from a recent investigation (RAIB, 2014) was the question of how to address booking on as fit for duty when staff may be fatigued.

As a direct result of the Llandovery investigation, RAIB undertook to review the state of science and practice regarding monitoring, reporting and mitigating the risk arising from fatigue for rail workers. We conducted a review of literature and practices in rail and other safety-critical industries, aimed at understanding barriers to reporting as fit for duty when possibly suffering from fatigue, as well as identifying and managing the accumulation of fatigue over the duration of a shift.

Fatigue management

Working hours

The classical approach to managing fatigue is in limitations on working hours. In the UK rail industry, this was manifest in the Hidden recommendations following the Clapham Junction accident in 1988, where the effect of long working hours was seen to be a factor that contributed to the accident. Working hour limits are still used in the UK and Canadian rail industries to manage fatigue.

Flight time limitations (FTLs) as a control measure are now prevalent across the aviation industry. The maritime industry is perhaps less mature in its awareness and management of fatigue; maritime watch patterns, coupled with pressures on resourcing levels, are particularly challenging for fatigue management. Currently, working hours limits are stated in some marine regulators' guidance notes.

However, many industries have now concluded that limitations on working hours are too simplistic for fatigue management, and are moving towards more scientific approaches. The logical extension of working hours limits is to design shift rosters that take account of human fatigue.

Rostering and fatigue modelling

The current trend in many industries is to assess fatigue risk in designing shift rosters. Initially, this involved simple rules such as forward rotation of shifts, or a clear night's sleep between late and early shifts. In the same vein, a RAIB investigation into an incident at Shap (RAIB, 2011) highlighted the risks of first night shifts. More recently, researchers in this area have produced a variety of biomathematical fatigue models, which attempt to predict fatigue and risk and are variously based on hours of work, rest times, time of day, and sleep / wake times. Such models are widely used, with domain-specific versions having been developed for the aviation, rail and marine sectors.

Whilst there is good evidence that an appropriately designed roster can reduce the risk of performance problems, the science base is yet to establish a clear relationship between fatigue and the types of risk or error. Moreover, the limitations of these fatigue models have been widely discussed, with the RAIB's report on the incident at Shap (referred to above) providing a detailed critique of several models. The fatigue scores only provide an approximation of the risk and fail to take into account variations in workload, individual differences or cumulative fatigue.

One of the main drawbacks with some of these models is that they are primarily based on working hours as defined on the roster – they do not account for actual sleep quantity or quality. Dawson and McCulloch (2005) pointed out that the fatigue 'clock' starts ticking from the moment of waking, not from the time of starting work. Their 'prior sleep and wake model' therefore uses a simple algorithm based on the amount of sleep obtained in the last 24 and 48 hours, and the length of time awake from waking up to the end of the shift.

Translating this into shift planning, there is growing realisation that rest (i.e., non-working) hours do not equate to hours asleep. This has led some organisations and regulators to introduce the notion of 'protected hours' – usually at times of peak anticipated fatigue, when the employee must be properly resting.

In light of the debate about the validity of biomathematical fatigue models, there is currently a movement away from using such quantitative thresholds. In any case, rostering practices should be part of a wider organisational fatigue risk management system (FRMS).

Fatigue risk management systems

Fatigue risk management systems (FRMS) typically cover roster design, the identification, reporting and monitoring of fatigue, the development of policy and procedure, training and education, and evaluation of the overall system. Guidance documents are available from the regulators in the rail, aviation and marine industries, although the rate of adoption by operators across these sectors is variable.

A key characteristic of FRMS is the guidance offered to both employer and employee in managing fatigue – implying a commitment on both sides. While the employer should offer rest breaks and suitable environments to take rest, the employee can take steps to adapt their lifestyle and sleep hygiene to ensure that they are properly rested. In order for this to be successful and to overcome ingrained attitudes towards fatigue, employees need to be engaged in the process.

Education and engagement

Most FRMS guidance documents include some element of fatigue awareness training, with different syllabi targeted at front-line crew as well as management levels. Broadly, such training covers the causes and effects of fatigue, signs and symptoms, and strategies for coping with fatigue – which can include both lifestyle guidance as well as short-term countermeasures while on shift. In addition, managers receive training on rostering and crewing.

The success of a fatigue management programme depends not only on education, but also on employee engagement. Successful engagement and considered training of the workforce in fatigue management can improve the quality and quantity of fatigue reporting – a key safety performance indicator. In turn, this is dependent on the organisational culture.

Culture and reporting

Although the FRMS should mitigate against fatigue, there will still be occasions when staff are fatigued at work. These employees need to feel able to raise the issue before it becomes a problem, just as they would for any other issue with health and well-being. Organisational culture can be a barrier or a facilitator towards reporting as fit (or otherwise) for duty. On the one hand, employees might feel reticent to declare themselves as unfit due to fatigue, perhaps because of peer or management influences, or the thought that it will increase workload for their colleagues. On the other, a 'just culture' would recognise the fatigue hazard and make it acceptable to report as unfit.

In such a just culture, suitable allowances are made for controllable fatigue – that is, fatigue due to working hours. The 'just' response is that these do not trigger an absence management process, as opposed to non-roster-related fatigue, which would be treated in the same way as a sickness absence. This approach also emphasises the joint responsibility of employer and employee in managing fatigue.

Regardless of the source of fatigue, employers should recognise the safety risk and have contingencies in place to cope with staff who declare themselves as fatigued. These might include having spare staff available or providing adequate opportunities and facilities for rest while on duty. Line managers play a key role in this process, not just in managing the operational response, but also in detecting and monitoring fatigue. However, most, if not all of these measures imply additional costs.

In terms of monitoring, the complement to fitness for duty reporting is in reporting incidences of fatigue on duty, which is used to monitor trends in fatigue across different roster patterns. Again, a just culture underpins both the quality and quantity of reporting, as employees need to know that they can report fatigue without facing any undue consequences for their jobs. However, it is worth noting that this may be open to abuse (i.e., using fatigue to condone poor performance).

Organisations vary in the effectiveness of their reporting schemes. As well as culture, awareness and training, one key driver seems to be ease of use in the reporting process. Whether an electronic or paper reporting system, forms need to be quick and easy to complete, and pre-populated where possible.

As well as reactive reporting, many guidelines encourage proactive monitoring of fatigue through fatigue surveys, either on a longitudinal basis or as a snapshot at a point in time. Whilst this can bring its own problems in processing large volumes of data, it is seen as good practice in identifying trends in fatigue.

Any fatigue monitoring programme is dependent on the ability to recognise the signs of fatigue, both on the part of the employer and the employee. But this is notoriously difficult – people are bad at judging their own fatigue state, let alone evaluating the claims of others. Recently, a number of techniques and technologies have emerged with the aim of making this judgement more objective.

Fatigue monitoring

Assessment

Given the limitations of biomathematical models discussed above, efforts have been directed towards a more objective or quantitative means of assessing fatigue risk in the context of reporting as fit for duty. Whilst several options have been explored, at present there is a need for further scientific validation of these metrics.

Probably the easiest measure to implement is a simple subjective rating scale, of which numerous examples exist (see Shahid et al., 2012). Subjective ratings have been used in aviation, military and marine environments, and show good associations with sleep history and circadian phase (Gander et al., 2015). Some mobile applications have even been developed for self-assessment of fatigue, using sleep/wake times and shift duration to calculate a fatigue score based on an underlying biomathematical model.

The problem with self-reporting is that it is vulnerable to unconscious bias, as well as the unreliability of judging one's own fatigue state. A more objective measure of sleep quantity is an activity monitor – known as an 'actigraph' or 'actiwatch'. This is a simple wristwatch-type device that registers movement and records the data for later download to computer. The data can then be used to infer sleep / wake history. Actigraphs have been used in several studies on sleep and fatigue, particularly in aviation (e.g., Roach et al., 2010), and can also be used to monitor fatigue trends of roster patterns over time.

Simple performance tasks have been used as a screening method to determine fatigue, using vigilance as a surrogate for fatigue. The psychomotor vigilance task (PVT) measures reaction time to visual stimuli, and can be presented on a handheld device such as a smartphone or tablet (some mobile apps are available that include a PVT). Typically, the PVT lasts for 5-10 minutes, but Basner and Rubinstein (2011) suggest that even a 3-minute version can be a reliable predictor of fatigue.

Whilst self-reports, actigraphs and the PVT are relatively easy to implement, the gold standard in fatigue measurement involves physiological techniques. These could include electrical activity in the brain (EEG), biomarkers (eg, in saliva, blood or urine), heart rate data or involuntary ocular reflexes to visual stimuli. Such metrics are more intrusive and need specialist equipment and expertise, making them largely impractical in an operational environment. A non-invasive technique is theoretically possible but still some years away from the current state of technology.

Monitoring

The ease of use of metrics such as the subjective rating scales and the PVT means that they can be applied for an on-the-spot measurement of fatigue during a task. For example, the aviation industry has trialled the use of the Samn-Perelli scale at key points in the flight, presented and recorded via the flight management system (Powell et al., 2011). Similarly, various mobile apps for reaction time testing can be used by flight crew at appropriate times while on duty. However, care must be taken that using the techniques in this way does not interfere with the primary task.

Less intrusive metrics fall into behavioural or physiological categories. In the automotive industry, systems are now available that monitor driving performance and provide an alert message to the driver if it deems that the car is being driven in an uncontrolled manner representative of fatigue or distraction. Although there is scientific evidence relating some of these driving performance variables to fatigue, establishing benchmarks and thresholds against natural variability in driving pose challenges to the validity and reliability of these systems. While similar studies have been conducted in aviation to identify performance correlates of fatigue (e.g., Srivastara and Barton, 2010), it remains to be seen whether such a system could be applied to train driver performance.

Other vehicle manufacturers use eye-tracking technology to monitor fatigue, and this is currently an area with much development work going on across several industries. These work well in relatively controlled environments (for instance, they have been successfully applied in studies of HGV drivers), but have limitations under certain lighting conditions or where the task involves continual head movements (such as with flight crew; Mallis et al., 2004).

Several papers (e.g., Greely et al., 2013) have examined the use of speech analysis to detect fatigue. The research has found a good association between voice production and measures of fatigue. However, the quality of speech recognition is heavily dependent on ambient noise in the environment.

Again, EEG and other physiological metrics offer the most valid measures of fatigue – but suffer from practical limitations in the operational environment. One project (RSSB, 2002) reviewed a number of such devices, including head and eye tracking, reaction times, EEG and skin conductance, concluding that the latter was the most feasible at the time. Head or eye tracking systems also had potential for railway applications – but at the time, the technology was not mature enough.

Many commentators still feel the same way about the current state of the art. Two recent review papers (Caldwell et al., 2009; Dawson et al., 2014) on these technologies found little evidence of scientific validation, concluding that none were suitable replacements for regulatory limitations. In any case, there remains the dilemma of how to respond if fatigue is detected.

Countermeasures

The only real cure for fatigue is sleep; this is why FRMS documents are without exception based on a premise of prevention rather than cure. However, it is acknowledged that, even with the best intentions, there will be situations when fatigue sets in while at work. Thus, whilst countermeasures are very much seen as a last resort, various strategies are used to enable a worker to get through their shift.

Given the restorative properties of sleep, a short nap can offer respite and is encouraged for motorists, long haul flight crew and ship crews, but is a less practicable solution in industries such as rail where there is little spare capacity and less use of multi-crew operations. The only opportunity for napping then comes on rest breaks, and some organisations do provide appropriate facilities in crew rooms. However, some studies caution that napping only relieves fatigue, and does not necessarily reduce accident risk. Alternatively, some recommend light physical activity during breaks as a means of raising energy levels.

If napping is the most widely recommended countermeasure, the next most popular solution is taking caffeine. A 150 mg dose of caffeine (roughly equivalent to a cup of filter coffee) can raise alertness within 15-20 minutes (and is hence sometimes used in conjunction with a nap), and its effects can last from one to two hours, (Reyner and Horne, 1997). Individual differences in tolerance cause a problem in determining an appropriate dose, though, leading some (e.g., Caldwell et al., 2009) to recommend 'tactical' caffeine use – only using it when it is really needed (such as at times of peak fatigue or before driving home from night duty). As an alternative to drinking coffee, several varieties of caffeine chewing gum are now available, which take effect more rapidly and have been demonstrated to improve performance in military studies. However, Moore-Ede et al. (2009) suggested that the cost of any temporary improvement in alertness may be repaid as excessive sleepiness after the effects have worn off. Moreover, any form of caffeine could affect sleep quality if taken within a few hours of trying to sleep, depending on individual differences.

A more novel means of boosting alertness is the use of light, which has been used in the past to shift circadian rhythms (ie, jet lag) but may also have a short-term effect on fatigue (Caldwell et al., 2009). Whilst further validation work is needed, some studies (see e.g., Beaven and Ekström, 2013; Lockley et al., 2006) suggest that relatively low levels of short wavelength (blue) light have more beneficial effects on performance than strong coffee. In an operational context, the light could be integrated into instrument panels or room / cab lighting. As with caffeine, though, there may be habituation or side-effects for later sleep quality; more research is necessary in this area.

The fatigue detection and monitoring technologies discussed earlier can be used as a basis for fatigue alarms. For example, some of the actigraph products include an inactivity alarm if movement is not detected within a certain period of time. But this assumes that a vigilant worker is constantly moving – which may not be the case in, for instance, a train cab. Vigilance alarms can also be used to mitigate fatigue; in the marine industry, watchkeeping alarms are encouraged for this purpose, but in many cases these might not be fitted or used by the crew (MAIB, 2004).

Finally, the effects of fatigue can be mitigated by restructuring the task or reassigning an employee to lighter or less safety-critical duties, where possible. For instance, a crew on a night shift could carry out the most safety-critical tasks early in the shift, before the so-called 'window of circadian low' in the early hours of the morning (Hobbs et al., 2011). Similarly, knowledge-based tasks – which are more susceptible to fatigue than skill-based tasks – could be rescheduled to avoid likely fatigue. Monotonous tasks, such as long-distance driving, have been known to introduce passive fatigue; the introduction of a secondary alertness-maintaining task can offset this. But all of these solutions depend on whether there is suitable flexibility in the task and the industry to allow such restructuring. In the rail industry, only some of these solutions will be considered as feasible and acceptable.

Conclusions and recommendations

Effective management of fatigue should be preventative, controlled through a properly implemented FRMS. Reactive fatigue detection and monitoring can form part of a FRMS, but should not replace the need for proactive fatigue management programmes. Countermeasures should only be used as a last resort.

Nevertheless, individual attitudes and organisational cultures surrounding fatigue constitute barriers to fatigue management, as there may be perceptions that colleagues will be let down, the job needs to be done, or – worse – that working through fatigue is a 'badge of honour'. In these cases, techniques and technologies for assessing and counteracting fatigue can play a role.

In terms of potential lessons for the rail industry, this paper argues that:

- rosters based on fatigue models often do not account for time spent sleeping (ie, allowing for time taken to travel and/or fall asleep), instead basing their assumptions on rest hours;
- the notion of 'protected hours' has been used elsewhere to ensure core minimum hours of sleeping in roster schedules;
- several techniques and technologies for detecting fatigue have been tested in other industries, but are subject to further validation; and
- a number of potential countermeasures are available that can offset fatigue, but more research is needed on potential side effects and feasibility assessments.

The overall message with regards to fatigue detection technologies is that the current state of the art does not offer a single, reliable and practical solution. Some reports recommend using a battery of measures, combining physiological or behavioural state with performance metrics; in practice, the choice of methods will be constrained by the operational context.

A key question surrounding the use of fatigue detection techniques is the threshold at which an alert is triggered – both in terms of defining the threshold and in using it. Ideally, the threshold should correlate with performance decrements, but research shows that fatigue dissociates from accident risk. Then there is the problem of how to use a threshold if it is breached. One of the problems with biomathematical models is that thresholds came to be seen as targets – given a quantitative limit, rosters were developed to that limit rather than in the spirit of the model to minimise fatigue. There is an associated risk that a user of a fatigue detection system (such as a train driver) might come to rely on it and be tempted to continue when they are actually fatigued. Moreover, if a train driver succumbs to fatigue during a journey, it is not especially practical to take a break for a nap.

Thus the countermeasures relevant to the rail industry are limited to the use of caffeine or blue light – both of which need further work in order to fully understand their effects. However, fatigue countermeasures need not be restricted to the human operator. Many road vehicle manufacturers are implementing accident avoidance technologies to prevent or mitigate the consequences of an accident. Such technology already exists in various forms on the railway; evaluating its relevance as a fatigue countermeasure could offer another route forward.

References

Åkerstedt, T. 1998, Shift work and disturbed sleep/wakefulness, *Sleep Medicine Reviews*, 2(2), 117-128

Basner, M. and Rubinstein, J. 2011, Fitness for duty: a 3 minute version of the Psychomotor Vigilance Test predicts fatigue related declines in luggage screening performance, *Journal of Occupational and Environmental Medicine*, 53(10), 1146–1154

Beaven, C.M., and Ekström, J. 2013, A comparison of blue light and caffeine effects on cognitive function and alertness in humans, *PLoS ONE*, 8(10), e76707

Caldwell, J.A., Mallis, M. M., Caldwell, J. L., Paul, M. A., Miller, J. C. and Neri, D. F. 2009, Fatigue countermeasures in aviation, *Aviation, Space and Environmental Medicine*, 80(1), 29-59

Dawson, D. and McCulloch, K. 2005, Managing fatigue: It's about sleep, *Sleep Medicine Reviews*, 9, 365-380

Dawson, D., Searle, A.K. and Paterson, J.L. 2014, Look before you (s)leep: Evaluating the use of fatigue detection technologies within a fatigue risk management system for the road transport industry, *Sleep Medicine Reviews*, 18(2), 141-152

Gander, P.H., Mulrine, H.M., van den Berg, M.J., Smith, A.A.T., Signal, T.L., Wu, L.J., and Belenky, G. 2015, Effects of sleep/wake history and circadian phase on proposed pilot fatigue safety performance indicators, *Journal of Sleep Research*, 24(1), 110-119

Greely, H.P., Roma, P.G., Mallis, M.M., Hursh, S.R., Mead, A.M. and Nesthus, T.E. 2013, *Field study evaluation of cepstrum coefficient speech analysis for* *fatigue in aviation cabin crew*. Report no. DOT/FAA/AM-13/19. (Federal Aviation Administration, Washington, DC)

Härmä, M., Sallinen, M., Ranta, R., Mutanen, P. and Műller, K. 2002, The effect of an irregular shift system on sleepiness at work in train drivers and railway traffic controllers, *Journal of Sleep Research*, 11(2), 141-151

Hobbs, A., Avers, K.B. and Hiles, J.J. 2011, *Fatigue risk management in aviation maintenance: current best practices and potential future countermeasures*. Report no. DOT/FAA/AM-11/10. (Federal Aviation Administration, Washington, DC)

Ingre, M., Van Leeuwen, W., Klemets, T., Ullvetter, C., Hough, S., Kecklund, G., Karlsson, D. and Åkerstedt, T. 2014, Validating and extending the three process model of alertness in airline operations, *PLoS ONE*, 9(10), e108679

Lockley, S.W., Evans, E.E., Scheer, F.A.J.L., Brainard, G.C. and Czeisler, C.A. 2006, Short-wavelength sensitivity for the direct effects of light on alertness, vigilance, and the waking electroencephalogram in humans, *SLEEP*, 29(2), 161-168

MAIB, 2004, *Bridge watchkeeping safety study*, Safety Study 1/2004 (Marine Accident Investigation Branch, Southampton)

Mallis, M., Neri, D.F., Colletti, L.M., Oyung, R.L., Reduta, D.D., and Van Dongen, H. 2004, Feasilbility of an automated drowsiness monitoring device on the flight deck, *Sleep*, 27. A167-8

Moore-Ede, M., Trutschel, U., Guttkuhn, R., Aguirre, A. and Heitmann, A. 2009, Fatigue countermeasure rebound: temporary alertness gain from caffeinated chewing gum repaid as excessive sleepiness after countermeasure cessation. *Proceedings of the International Conference on Fatigue Management in Transportation Operations*, Boston MA, March 24-26

Powell, D.M., Spencer, M.B. and Petrie, K.J. 2011, Automated collection of fatigue ratings at the top of descent: a practical commercial airline tool. *Aviation, Space and Environmental Medicine*, 82(11), 1037-41

RAIB, 2011, *Uncontrolled freight train run-back between Shap and Tebay, Cumbria, 17 August 2010*. Report 15/2011 (Rail Accident Investigation Branch, Derby)

RAIB, 2014, *Near-miss at Llandovery level crossing, Carmarthenshire, 6 August 2013*. Report 11/2014 (Rail Accident Investigation Branch, Derby)

Reyner, L. and Horne, J. 1997, Suppression of sleepiness in drivers: combination of caffeine with a short nap, *Psychophysiology*, 34(6), 721-725

Roach, G.D., Darwent, D. and Dawson, D. 2010, How well do pilots sleep during long-haul flights? *Ergonomics*, 53(9). 1072-75

RSSB, 2002, Driver vigilance devices – systems review. T024 report (RSSB, London)

RSSB, 2012, *Managing Fatigue – A Good Practice Guide*. Document no RS/504, issue 1 (RSSB, London)

Shahid, A, Wilkinson, K., Marcu, S. and Shapiro, C.M. 2012, eds., *STOP, THAT and One Hundred Other Sleep Scales*, (Springer, New York)