Sleepiness On The Flight Deck: Reported Rates Of Occurrence And Predicted Fatigue Risk Exposure Associated With UK Airline Pilot Work Schedules

Abstract

Survey and field studies conducted with commercial airline pilots suggest that in-flight sleepiness and related involuntary sleep phenomena are experienced by pilots during their duties. However, for methodological, practical and commercial reasons, there is a lack of publicly available research data of per-flight hour rates of sleepiness experienced by pilots or predicted fatigue risk rates associated with pilots' hours of work. This empirical field study sought to address this gap by collecting self-reported sleepiness/alertness ratings from pilots from 18 different UK airlines via a mobile phone app over the period of August 2017. In tandem, predicted sleepiness levels and sleep lengths associated with participants' flown rosters were investigated using biomathematical fatigue modelling. Findings indicated that a quarter of all flying duty periods are predicted to be preceded by a main sleep opportunity that is shorter than six hours, whilst 10% of all flying hours are associated with elevated fatigue risk levels. Pilots reported 7.3 reports of involuntary sleep on the flight deck per 1000 flying hours, which represents a rate far greater than that previously reported to the regulator. By comparison, the rates of predicted and recorded fatigue-related incapacitations greatly exceeded the target rate of medical incapacitation permissible under the medical incapacitation safety standard for commercial aviation of less than one occurrence per 1,000,000 hours.

Tags: fatigue; sleepiness; incapacitation risk; safety; aviation

1. Introduction

The hazard of operator fatigue has long been recognised as a significant potential risk to safe pilot performance within commercial aviation operations (National Transportation Safety Board, 2019; Caldwell, 2012, 2005; Wilson et al., 2007). Pilot fatigue has been implicated as a contributory factor to aircraft crashes and serious incidents on a number of occasions (Drury et al., 2012; NTSB Aircraft Accident Reports, 1993: AAR-04/94; 2009: AAR10-01; Rosekind et al., 2000; Swiss Aircraft Accident Investigation Bureau (SAAIB) Report, 2001: No.1793). In addition, fatigue has frequently been suggested to have affected both operating flight crew, highlighting the vulnerability of flight crew to simultaneous fatigue-related incapacitation risks (e.g. NTSB Reports: NTSB, 2017: NTSB/AIR-18/01, pp52-53; 2009: NTSB/AAR-10/01, pp106-107). Where pilot fatigue is implicated, investigators have typically used evidence-based inferences from sleep science principles, to determine whether sleep loss and extended wake circumstances are likely to have been caused by pilots' rostered duty hours or likely sleep-wake history preceding the duty. Reviews of commercial aviation crash reports have concluded that in at least 4-8% of crashes, fatigue is likely to have played a contributory role (Caldwell, 2005), and that duty time is linked with an increased likelihood of crash risk (Goode, 2003). However, some caution should be taken over the broader extrapolation of these rates, since the processes by which crashes or serious incidents have previously been categorised as 'fatigue-related' are not entirely known and may depend on available circumstantial evidence (Lyman & Orlady, 1981; Pouliquen et al., 2005).

Definition of pilot fatigue and sleepiness

The definition and use of the term 'pilot fatigue' within the aviation industry reflects a safety hazard concern over the impact that sleep, circadian and cognitive work load factors may have, on their own or in combination, on pilot performance and functioning in-flight. As set out by the International Civil Aviation Organisation (ICAO, 2012), ‘Pilot fatigue’ is functionally defined as: “A physiological state of reduced mental or physical performance capability resulting from sleep loss or extended wakefulness, circadian phase, or workload (mental and/or physical activity) that can impair a crew member’s alertness and ability to safely operate an aircraft or perform safety related duties.”. In practice, sleep and circadian drives have emerged as the key factors of research interest with regards to pilot fatigue, and the term is often used interchangeably with ‘sleepiness’ (the brain state associated with instability of wakefulness) and the increasing physiological propensity to fall asleep (Durmer & Dinges, 2005; Phillips, 2015). This is because the sleepiness state has received greater academic consensus on its definition, biological causes, measurement and associated performance decrements (Åkerstedt et al., 2014; Caldwell, 2012; Cheng & Drake, 2016; Horne & Reyner, 1999).

Sleep drives and biomathematical modelling

Among the many factors that influence sleepiness, some can be quantified reasonably well in terms of their effects on alertness and neurobehavioural functioning (Åkerstedt & Folkard, 1996; Dawson et al., 2011; McCauley et al., 2013). Experimental studies suggest that on many aspects of cognition, overall performance declines as a function of time spent awake, and this decline in performance is modulated by circadian rhythm (Durmer & Dinges, 2005; Goel et al., 2013). For the typical day, this means that subjective sleepiness and sleep-driven performance lapses are low across the first 16 hours of wakefulness, but then increase across the habitual night, peaking at around 26 hours awake (Åkerstedt & Wright, 2009). To date, the most useful model for predicting physiological sleepiness and the likely sleep-wake cycle in humans is the two-process model, that mathematically charts the interaction of sleep and circadian factors (Borbély, 1982; Borbély & Achermann, 1992). The homeostatic sleep drive is modelled as a pattern of increasing sleepiness with increasing periods of continual wakefulness, and the recovery of alertness during sleep. The circadian processes are described by a pair of sinusoidal waves, the circadian rhythm lasting twenty-four hours, and an ultradian rhythm lasting twelve hours. The aggregation of these processes produces an estimated level of sleepiness at any given moment. The overall effect of this mathematical profile is that humans are alert throughout most of the morning, afternoon and early evening (save for a small dip in the early afternoon), but that this alertness decreases quite rapidly as the night progresses, where the drive from both sleep and circadian factors is towards sleep (Basner et al., 2013). Although there may be some differences in attempts to mathematically chart this profile depending on e.g. presumed light availability (Shen et al., 2006) or chronotype (Kerkhof & Van Dongen, 1996), many established models of sleepiness, if not all, appear to be founded on this baseline formulation. While the two-process model is commonly used to make predictions of alertness and sleepiness, its original purpose was to make predictions regarding the timing and duration of sleep (Dawson et al., 2011). Given a specific pattern of work, the two-process model is capable of making predictions regarding the timing and duration of sleep that an average person would experience, to a reasonable degree of accuracy (Dorrian et al., 2012). Extending beyond this, several biomathematical models that predict sleepiness or fatigue levels during waking hours have been validated against performance in laboratory, driving, aviation and shift work settings (Åkerstedt et al., 2008; Dawson et al., 2011; Ingre et al., 2014; Kandelaars et al., 2005; Van Dongen, 2004). The use of biomathematical models to predict sleep opportunities and approximate on-duty sleepiness has hence become particularly important in work environments where operator fatigue risks are elevated by the intrinsic nature of different types of shift patterns.

Shift work and sleep drives

Like other forms of shift work, the timing and duration of pilot duty hours often come into conflict with homeostatic and circadian drive aspects of human functioning to cause elevated sleepiness levels during duty and truncated sleep at night. Hence many features of shift work - early start times, extended work periods, truncated recovery time periods between duties, night work through the window of circadian low, daytime sleep periods and day-to-night or night-to-day transitions across consecutive work periods – can act alone or in combination to increase sleep loss and on-duty sleepiness in pilots (Caldwell et al., 2009; Gander et al., 2014; Roach et al., 2012). In addition to these common shift work factors, commercial flight crews may also have rosters that cause circadian rhythm desynchrony (jet lag) from the crossing of multiple time zones, or face a number of other environmental factors or work pressures from the cognitive demands of the piloting role that may affect their individual fatigue levels (Caldwell, 2012). Against this context it has been suggested that real world investigations into elevated fatigue risks in safety-critical operators should begin by assessing scheduling practices for insufficient sleep opportunities afforded by work and extended time on duty (Dawson & McCulloch, 2005).

Despite widespread recognition that sleepiness degrades neurobehavioural performance in humans, and further operational evidence implicating ‘pilot fatigue’ as a risk factor in commercial flights, it has been difficult for practitioners to determine at which point this risk becomes unacceptable. Data from other transport and operational settingssuggests that there are elevated fatigue-related risks where continual wakefulness exceeds 16 hours, the sleep obtained prior to duty start is shorter than six hours, or the duty occurs during the individual’s usual sleep hours (National Research Council, 2011; Van Dongen et al., 2003; Williamson et al., 2011). However, in terms of commercial aviation specific safety risks, to our knowledge, there have been no systematic attempts to investigate the dose-dependent effects of sleep loss or circadian influences on multi-crew commercial flight performance. It is for this reason that within aviation, sleepiness risk rates associated with duty patterns need to be compared not only against the established laboratory evidence on neurobehavioural performance decline, but also against other parallel safety standards that govern the functional status of the human operator with respect to safe flight performance.

Assessing human factors risks in commercial aviation

In commercial aviation, a principle of designing and certificating aircraft for safe flight is that the various sub-systems of the aircraft (the engines, the electrical systems, the pilots, etc) should meet a quantified reliability standard (Zio et al., 2019). To minimise any risk of a ‘weak link’ in the chain, each sub-system should ideally meet a similar reliability standard. In many cases, in order to achieve this standard, safety critical components of the sub-system are at least duplicated, and sometimes triplicated or more (Tunstall-Pedoe, 1988). Hence, for the sub-system that sustains powered flight, where commercial aircraft have two engines, the design is such that each engine on its own can sustain flight. Similarly, where there are two pilots, in the event of one of the pilots becoming medically incapacitated, the other pilot on their own can continue the flight safely. Thus pilots are also considered to be part of the system that meets a stringent reliability standard, to prevent the sudden or subtle, partial or complete 'incapacitation' of the individual due to the effects of a medical condition or a physiological impairment that could represent a potential threat to flight safety (Australian Transport Safety Bureau, 2007).

The International Civil Aviation Organisation (ICAO) guidance on the medical incapacitation standard for pilots (of no more than 1 occurrence per 1,000,000 hours) is the current acceptable rate of risk for the likelihood of the break-down of optimal or safe performance of the operating flight crew where the cause is 'medically-driven' (ICAO, 2012; Mitchell & Evans, 2004) and represents a probabilistic standard concerning pilot safe functioning in-flight. Such an approach is intended to ensure that individuals who are granted a flight crew licence for commercial aviation activities represent a medically fit pilot population, at an acceptably low risk of likely in-flight performance impairment or incapacitation. The medical causes considered as likely to cause performance decrement that could represent a potential threat to flight safety range from sudden serious events such as heart attacks and epileptic fits, through to more subtle events, such as headaches, that still are capable of inducing considerable performance decrements (ICAO, 2006). Against this context, it should be noted that although there are European and national regulations on pilot duty hours (UK Civil Aviation Authority, 2019), these regulations do not attempt to quantify the acceptable level of fatigue-driven incapacitation risk in the same way as the medical incapacitation risk standard. However, in terms of consistency within safety systems, it follows that acceptable risk rates of flight crew performance decrements and incapacitations due to sleepiness need to be considered within the same risk framework as medical causes.

Indeed, sleep-related incapacitations affect sensory, cognitive, physical and behavioural functioning of crew and thus are functionally similar to a variety of medically-driven incapacitations. Neurobehavioural performance effects of sleepiness reported by pilots include increasing pressure to fall asleep, degraded alertness, errors of omission and commission, deterioration in judgement and decision making, worsened mood, and deteriorating flying skills (Bourgeois-Bougrine et al., 2003; Dinges & Kribbs, 1991; Petrie, Powell, & Broadbent, 2004; Petrilli et al., 2006; Rosekind et al., 2000; Samel, Wegmann, & Vejvoda, 1997). Field studies using objective electroencephalography (EEG) and electrooculography (EOG) techniques in flight have revealed that significant sleepiness and involuntary sleep events occur in the commercial aviation setting, and may occur without the pilots' awareness (Civil Aviation Authority Safety Regulation Group, 2003; Wright & McGown, 2001). Involuntary sleeps on the flight deck lasting from 20 seconds to longer than 10 minutes (Civil Aviation Authority Safety Regulation Group, 2003; Graeber, Rosekind, Connell, & Dinges, 1990) as well as periods of simultaneous sleepiness in both captain and co-pilot (Cabon et al.,1993), have also been reported in other studies using EEG recordings (Rosekind et al., 1994; Samel et al., 1997).

Whilst it may be argued that serious forms of medical incapacitation, such as cardiovascular events, are inherently more dangerous to individual flight crew due to the sudden and complete loss of function, it is likely that such medically-driven risks are independent between flight crew members. By contrast, sleep-driven risks are more likely to co-occur for flight crew undertaking the same duty patterns, and as such, may represent a greater threat to the overall safety of flight (Gander & Signal, 2008). Instances of involuntary sleep on the flight deck are widely reported among commercial airline pilots, with estimates ranging from 56% to 71% of commercial airline pilots having experienced it at some point in their careers (ComRes, 2013; Rosekind et al., 2000). In terms of simultaneous sleep-driven incapacitation, in one survey, almost a third of pilots who had reported having involuntarily fallen asleep on the flight deck also reported to have also woken up to find the other flying crew member had fallen asleep (ComRes, 2013). This real world and experimental evidence of the impact of sleepiness on neurobehavioural functioning, and the finding that pilots across Europe also regularly cite pilot fatigue as a significant threat to pilot performance and safety (European Cockpit Association, 2012), suggests there is little conceptual justification to have vastly different acceptable risk rates for sleep- and medical- related incapacitations. The medical incapacitation rate hence provides a useful aviation specific benchmark for appraising how sleepiness occurrences and predicted fatigue risk rates associated with pilot rosters may compare against other risks to operator safe performance.

Existing data on rates of occurrence & Purpose of study

Surveys have tried to identify the rates with which high levels of sleepiness or severe 'fatigue' are experienced in commercial aviation (e.g.. Houston et al., 2012; Petrie et al., 2004). However, it is difficult to determine comparable rates from the findings of such studies due to the different conceptualisations of fatigue used, or absence of similar time scales involved with respect to operational variables. Moreover, accurate rates have been difficult to collate from existing industry sources since there is a large degree of underreporting of fatigue from pilots via formal channels to the company and regulator (Confidential Incident Reporting Programme (CHIRP), 2017; Reis et al., 2013). Indeed, despite research data indicating high levels of involuntary sleep on the flight deck, in terms of regulatory data, just two reports of this occurrence were submitted to the UK aviation regulator between 1976 - 2013 (BBC Freedom of Information Act request to the CAA, F0001485, 2013). Hence, formal reports likely underrepresent the real world incidence levels. Prior research involving the monitoring of pilot fatigue or sleepiness levels during actual operations has also typically focussed on addressing a very specific set of operational issues for a particular airline or type of operation (e.g. Gander et al., 2013; Samel et al., 1997; Wright & McGown, 2001). As a result, the number and range of participating pilots from different airlines have understandably, tended to be somewhat restricted. Previous field and survey studies have not produced an overarching picture of both reported and predicted incidence rates of sleep related phenomena or high sleepiness levels occurring during flight, which are important metrics for understanding the risk exposure and assessment of the threat to safety across the aviation industry. The following study sought to overcome some of these constraints by investigating the predicted fatigue risks of actual flown pilot working hours alongside self-reported sleepiness ratings and involuntary sleep occurrences in UK Airline Pilots during August 2017. The aims were to 1) assess the severity of sleepiness levels experienced by pilots during short and long haul flights 2) assess reported involuntary sleep rates; and 3) to assess schedule-driven fatigue risk exposure and associated sleep opportunities approximated by biomathematical modelling of duty patterns.

Materials and Methods

2.1 Participants and recruitment

Participants were recruited via the British Airline Pilots' Association (BALPA) membership database, which represents approximately 85% of all UK commercial airline pilots. An invitation to take part in the study was sent to eligible BALPA full members. Exclusion criteria included pilots who were retired members or not currently employed, or to those who have unsubscribed from membership communications, who were not contacted. 294 pilots from 19 companies volunteered (3.12% of 9461 contacted members). Pearson's chi-squared test of independence revealed that the volunteer sample was not significantly different to the membership population in terms of company break down χ^2 (180) = 190, p=0.29), gender χ^2 (1) = 0.86, p=0.35 or age, t (312.11) = 1.93, p> 0.05. Further descriptive statistics of the pilot sample are provided in Table 2.

2.2 Measurements

2.2.1 Karolinska Sleepiness Scale (KSS)

The KSS is a nine-point one dimensional subjective scale, which is both sensitive to sleepiness fluctuations (Åkerstedt et al., 2014; Åkerstedt & Gillberg, 1990), and validated against performance and EEG variables (Kaida et al., 2006; Kecklund & Åkerstedt, 1993; Sagaspe et al., 2008). Rating on this scale can be done repeatedly and provides the individual with a simple measure of their sleepiness that reflects the psycho-physical state experienced in the last five minutes. The KSS is considered a reliable tool for use both within laboratory and field research for evaluating changing levels of sleepiness (Åkerstedt et al., 2014). Previous research has suggested that physiological markers of sleepiness such as long eyelid closures and slow eye movements start to occur at or above KSS 7 (Kaida et al., 2006). At KSS 8, these symptoms appear to substantially increase in frequency and occur for longer durations. There is also a marked increase in microsleeping risk, and top down attempts to stave of sleep are breaking down. Once KSS 9 is reached, sleep intrusions dominate EEG and EOG recordings (Åkerstedt & Wright, 2009). In plain terms, even motivated individuals find it difficult to stay awake at KSS 9, since they are actively 'fighting sleep'. KSS levels of 8-9 are also related to substantial increases in driving incident and crash risk (Åkerstedt et al., 2014; Reyner & Horne, 1998).

Table 1. Karolinska Sleepiness Scale (Åkerstedt & Gillberg, 1990)

1	Extremely alert
2	Very alert
3	Alert
4	Rather alert
5	Neither alert nor sleepy
6	Some signs of sleepiness
7	Sleepy but no effort to keep awake
8	Sleepy, some effort to keep awake
9	Very sleepy, great effort to keep awake, fighting sleep

2.2.2. The BALPA 2-way app: Subjective sleepiness ratings and involuntary sleep reports

The BALPA 2-way app is a free mobile app, which enables BALPA pilot members to communicate with BALPA via specific report forms. For this research, a section of the app was developed to enable pilots to rate their KSS levels during their waking hours. Following the selection of KSS level, the user is prompted to select the current date and time of rating in Coordinated Universal Time, (UTC). These times are subsequently converted into the correct local time zone depending on home base and time of year. For the present study, times were converted into British Summer Time (BST), since this was the local time zone for all participants who took part in this study. The form also provides the option for users to report any instances of involuntary sleep experienced on the flight deck during duty, and a free text option for pilots to add any additional narrative to the report, if desired. Reports within the app are automatically uploaded and transferred to a secure database when the user has an internet connection on their device. The date and time details that the user confirms at the point of rating are preserved, even if the report is only subsequently uploaded at a later stage, with each report recording the time stamp that the report was made, and when it was transmitted. The task of making a submission takes approximately 30 seconds to complete.

Pilots who agreed to participate in this study were asked to download the BALPA-2-way app onto their mobile device, and instructed to submit six or more KSS ratings both during flying and non-flying duty days, during August 2017. Participants were asked to rate at all levels of alertness (i.e. to submit KSS ratings not only when they were feeling sleepy, but also when they were feeling more alert). Participants were encouraged to spread their ratings across their hours of wakefulness, although exact timing of the ratings were necessarily determined by the individual pilot, depending on their waking and working hours. In addition to the KSS rating, participants were also instructed to indicate any instances of involuntary sleep during flying duty days, where they and/or their accompanying flight crew had involuntarily fallen asleep at any point during flight. Participants were told that there was an optional free text box to add any operational or additional sleepiness details.

2.2.3 Demographics

Participants were asked to provide demographic details (age, sex) as well as flight related details (role; captain or first officer, flight experience in years) and additional sleep-related information (commute time and type, habitual caffeine use and chronotype) via two online questionnaires at the start of the study period. For the assessment of chronotype, participants were asked to classify themselves as one of the following “very early”, “early”, “neutral”, “late”, “very late” (“One hears about morning and evening types of people, which one of these types do you consider yourself to be?” (Horne & Ostberg, 1976).

2.2.4 Roster information and modelling assumptions

At the end of the August study period month, participants were asked to send in their flown rosters (achieved work shift patterns) for the months of July and August. The July month was retrieved to enable more accurate biomathematical fatigue predictions to be made for the first week of August. This roster information (including specific flights, duty periods, airport destinations) was then inputted into the Sleep-wake Predictor (SWP) (SWP©: version 3.12; Åkerstedt et al., 2008) biomathematical model, for the computation of KSS predictions during work and waking hours, and prediction of assumed sleep lengths and timings. Within the range of commercially available models, we chose the Sleep wake predictor model (SWP) for research purposes as the full algorithms underpinning the modelling of circadian and homeostatic processes have been published (Åkerstedt et al., 2008). Rosters were entered into the model in the local time zone of home base, which for all participants was British Summer Time (BST). Key schedule information such as positioning duties (where pilots are travelling to their next flight within the duty period); standby duties (where pilots need to be prepared to be called for duty) and flying duties were extracted from the rosters. All other forms of activity (e.g. flight simulator training, ground training) were categorised as ground duties. Exact check-in and debrief duty times were inputted into SWP when this information was available in the rosters. For the minority of rosters which contained some, but not all of these full details, it was assumed that check-in prior to the flight time on rosters took place 1 hour prior (as the industry standard), and debrief period (the time from flight end till the end of duty period) was 30 minutes. Commute time was assumed to be 60 minutes as this is the default assumption in SWP. In addition, the majority of volunteers (61%) indicated that their usual commute duration was between 30-90 minutes. For long haul routes, a number of variables including company, time of day, length of route and destination determine whether or not the number of flight crew is two or more. For example, in some companies three pilots are provided for single flights over 9-10 hours, whereas in others this is not the case. Given the large number of variables which differ between long haul participants' rosters, a decision was made to favour fewer assumptions and thus not apportion additional sleep opportunities for long haul rosters, in the case that there may be additional flight crew. In the present study, it is estimated that this assumption could affect 344.58 flying hours, which was 5% of the entire flying hour sample. This estimation is derived by calculating the number of flight hours in the data where the pilots' roster indicated that they had been flying for eight or more hours continuously. Flights shorter than eight hours are rarely afforded an additional pilot and the possibility of in-flight rest. The allocation of an additional pilot beyond eight hours is subject to considerable variation, and the achievement of sleep during any rest period during flight is not guaranteed and may also vary widely. As such, it was reasoned that any attempt to assume availability of in-flight rest or voluntary sleep would be a greater source of imprecision, affecting a greater proportion of the sample, than the decision not to make such an assumption. Four helicopter pilots also opted into the study, with two providing schedule and rating information, but due to their small number and inherent difference in operation from airline operations, these participants' data were excluded from the analyses in this paper.

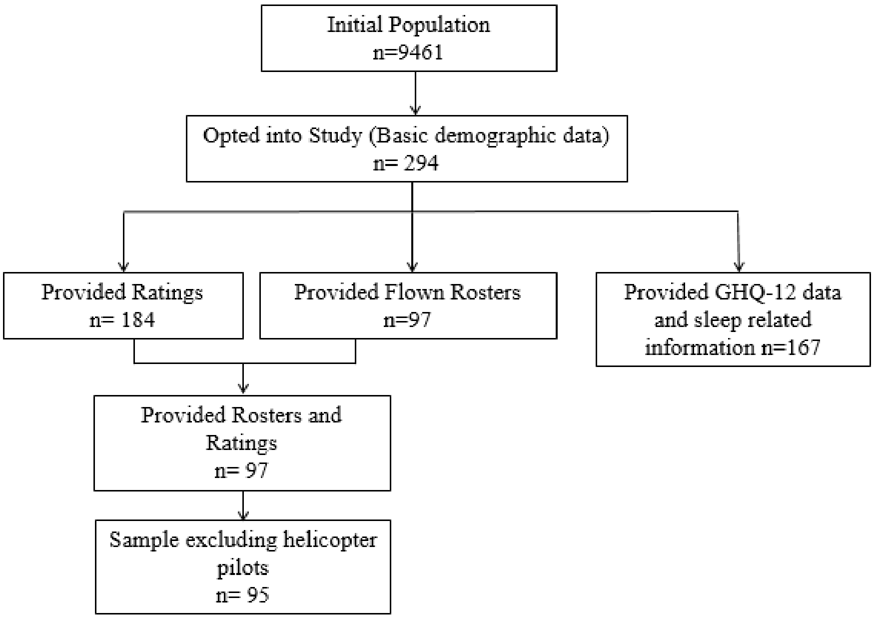

Figure 1. Flow chart indicating participant numbers for the different categories of data in the field study. The roster and rating data are the primary focus of this paper.

2.3 Ethics Statement

The study was approved by the University College London Research Ethics Committee, reference 8015/001. Participation was voluntary and informed consent was obtained by requiring prospective pilots to opt-in to the study via email. Participants did not receive any payment or reward for their time or effort.

3. Results

3.1 Demographics of study sample

Table 2 shows the demographic details of the pilot participants who opted into the study, and those participants who submitted both KSS ratings and complete rosters during the 1 month period.

Table 2. Demographic descriptive statistics of pilot participants

		Opt in n= 294		**Provided Rosters & Ratings n= 95**	
Sex					
	Male	272	92.5%	85	89.5%
	Female	22	7.5%	10	10.5%
Age					
	Years (mean ± s.d.)	42.1 ± 9.7		42.0 ± 10.3	
Flight experience		**n= 175**			
	Years (mean ± s.d.)	15.2 ± 9.0		15.6 ± 8.8	
Role		**n= 292**			
	Captain	154	52.7%	47	50%
	First Officer	135	46.2%	47	49%
	Other	3	1%	1	1%
Operation type					
	Long haul	n/a	n/a	23	24.2%
	Short haul	n/a	n/a	72	75.8%

Table values are provided to one decimal place. n/a refers to data that were not available; s.d. refers to standard deviation. Of participants that opted in, some provided incomplete data, so subset participant numbers are listed for flight experience and role.

3.2 Description of pilot working hours

Within commercial pilot work schedules, the term ‘Flying hours’ relates to the period of time between the aircraft being off blocks (when the aircraft becomes free to move) and on blocks (where the aircraft is restrained from moving). ‘Flight duty period’ (FDP) refers to a duty which includes flying time, the turnaround time, and one hour pre-flight preparation. A ‘Duty period’ may relate to a period of work which includes flying and/or non flying duties; where the duty period includes an FDP, the period will additionally include the debrief time at the end of a flight, which is approximately 30 minutes for most airlines. Typically long haul routes only include one sector, whereas short haul routes range from two to six sectors. Long haul operations are not normally followed by an immediate flight the next day, and so consecutive duty starts in this study almost exclusively refer to short haul operations. Table 3 shows flight and duty information from pilot’s work schedules. For the purposes of this study, ‘Super early start’ duties were defined as check in times between 00:00- 06:00 BST; ‘Early start’ duties referred to check in times between 06:01- 09:00 BST; ‘Late finish’ duties referred :to duty end times after 00:00 BST. Window of circadian low or ‘WOCL’ duties referred to duty starts between 00:00-06:00 BST or duty finishes between 0000-0800.

Table 3. Flight and duty information from pilot’s work schedules

Flight and duty information	Long-Haul	Short-Haul	Overall
Flight Period			
Average Flight length (hours) (mean, s.d.)	9.4 ± 2.6	2.2 ± 1.1	
Number of Flights	180	2414	2594
Number of Flight Duty periods	172	932	1104
Flying Hours (total)	1692.6	5294.1	6986.7
Duty Period			
Average Duty Length (hours) (mean, s.d.)	11.2 ± 3.1	8.5 ± 3.2	
Number of Duty Hours	2123.6	9374.3	11497.9
Number of Duty periods	190	1105	1295
Super Early Start Duties	43 (22.6%)	267 (24.2%)	310 (23.9%)
Early Start Duties	56 (29.5%)	533 (48.2%)	610 (47.1%)
Late Finish Duties	64 (33.7%)	115 (10.4%)	197 (15.2%)
WOCL Duties	106 (55.8%)	379 (34.3%)	485 (37.5%)
Consecutive duties starts before 06.00			
2 x Consecutive Super Early	1 (0.5%)	118 (10.7%)	119 (9.2%)
3 x Consecutive Super Early	0 (0%)	51 (4.6%)	51 (3.9%)
Consecutive duties starts before 0900			
2 x Consecutive Early	1 (0.5%)	326 (29.5%)	327 (25.3%)
3 x Consecutive Early	0 (0%)	188 (17%)	188 (14.5%)
Consecutive duties with late finishes			
2 x Consecutive Late Finish	1 (0.5%)	27 (2.4%)	28 (0.1%)
3 x Consecutive Late Finish	0 (0%)	3 (0.3%)	3 (0.2%)
Duty start times shifted (3 or more hours) relative to previous duty start times	32 (16.8%)	178 (16.1%)	210 (16.2%)

Within our sample of 1295 duty periods, duties that started or ended during the known WOCL periods made up over a third of all duties, with 310 (24%) having check in times before 0600 and 197 (15%) with duty end times after 0000. For both long and short haul pilots, over a quarter of duty start times in this study were shifted 3 or more hours relative to previous duty start times.

Table 4 shows the biomathematical predictions of main sleep opportunities prior to flying duties. ‘FDP’ relates specifically to a duty period which includes flying duties. Sleep periods of less than three hours duration (typically pre-flight ‘nap’ opportunities ahead of late duties) were excluded for this table.

Table 4. Biomathematical Predictions of main sleep periods prior to flying duties

Flying Duty Periods (FDPs)	Long Haul		Short Haul		Overall	
Predicted sleep length (mean ± s.d.)	7.4 ± 1.0		6.8 ± 1.2		6.9 ± 1.2	
Median predicted sleep length (IQR)	7.8 (7.5 - 7.8)		7 (5.8 - 7.8)		7.33 (6.0 - 7.83)	
Predicted sleep period <7 hours	30	17.4%	441	47.3%	471	42.7%
Predicted sleep period <6 hours	16	9.3%	259	27.8	275	24.9%
Preceded by three consecutive sleep periods<7 hours	0	0%	105	11.3%	105	9.5%
Preceded by three consecutive sleep periods <6 hours	0	0%	20	2.1%	20	1.8%

Welch's two-tailed t-test revealed that there was a significant difference between the mean predicted sleep lengths prior to flying duties between short and long haul work patterns t = 7.73, df = 281.4, p< 0.001, and examination of the interquartile ranges reveals the greater spread of predicted pre-flight sleep opportunities in short haul (5.8-7.8 hours) compared with long haul operations (7.5-7.8) hours. When considering the main sleep opportunities prior to duty, a quarter (24.9%) of flying duty periods were predicted to be preceded by a sleep period of less than six hours, extending up to 42.7% predicted to be preceded by a sleep period of less than seven hours. Since there were greater numbers of short haul pilots who both opted into the study and provided full roster information (SH n= 72; LH n= 23), the overall proportion reflects the greater abundance of short haul duties within the analysis. With respect to repeated shortened sleep periods, over 10% of short haul flying duty periods were predicted to be preceded by three consecutive sleep periods of less than seven hours, and 2.1% less than six hours.

Within our dataset a small proportion of flying duty periods (32 out of the study sample of 1104 FDPs) were associated with a pre-flight sleep period of less than three hours. Most of this subset of predicted sleep opportunities less than three hours (28 FDPs) related to long haul duties where the model predicts an additional day sleep period prior to a late departure flight. For the purposes of establishing the mean predicted sleep length prior to flight across all duties, such sleep opportunities were assumed to be a pre-duty nap and hence eliminated from the analysis in Table 3, since they did not represent the main sleep opportunity prior to flight and would disproportionately reduce the mean long haul pre-flight sleep lengths. However, for all following analyses pertaining to biomathematical model predictions concerning on-duty alertness, these predicted nap opportunities were retained.

3.3 Biomathematical Predictions of Karolinska Sleepiness Score (KSS) levels at Duty points

Table 5 shows that the mean predicted sleepiness scores at key operational points (duty start, and duty end) were overall not indicative of severe KSS sleepiness scores (predicted KSS 7 or above), in long haul or short haul flights in this study, although the standard deviation values indicate a reasonably large degree of variation around these means, particularly in long haul rosters.

Table 5. Predicted KSS Scores at operational duty points

	Predicted KSS level (mean, s.d.)	
Operation Point	**Long-Haul**	**Short-Haul**
Duty Start	5.06 ± 1.51	4.53 ± 0.85
Last Landing	6.57 ± 1.63	4.94 ± 1.05
Duty End	6.68 ± 1.68	4.96 ± 1.11

Duty operation points relate to times retrieved from participants' submitted flown work schedules.

As may be viewed in Table 6, elevated KSS levels were predicted to occur during approximately 10% of flying hours, within which 225.7 hours (3%) were associated with predicted fatigue levels of KSS 8 or above. The majority of flying hours associated with elevated KSS predictions were from long haul schedules. 12% of flying hours in the sample were associated with continual hours of wakefulness in excess of 16 hours. On average short haul pilots were likely to have been awake for 11.56 hours at last landing, (s.d. = 3.76), compared with long haul pilots with an average 18.31 hours, with a large deviation around the mean (s.d. = 5.6).

Table 6. Flying hours associated with KSS predictions in excess of 7 and continual hours of wakefulness in excess of 16 hours

	KSS Prediction		**Continual Hours Wakefulness**		
	KSS >=7	KSS >= 8	>=16 hours	>=17 hours	>=18 hours
Long Haul flying hours	591.7	208.8	647.0	558.2	465.0
% long haul flying hours	35.0	12.3	38.2	33.0	27.5
Short Haul flying hours	95.0	16.8	201.7	101.7	45.5
% short haul flying hours	1.8	0.3	3.8	1.9	0.9
Total	686.7	225.7	848.7	659.8	510.5
% of total flying hours in sample	**9.8%**	**3.2%**	**12.2%**	**9.4%**	**7.3%**

3.4 Karolinska Score Ratings

There were 8291 ratings provided by participants during the study period via the 2-way-App, and 5382 ratings from the 97 people who submitted their full rosters, with this subset of participants submitting an average of 1.8 ratings per day. Figures 2a and 2b show the relative distribution of ratings between operation type and time of day. The majority of ratings clustered around the middle values of the scale, suggesting that as instructed, in their waking hours pilots were providing ratings at all states of alertness/sleepiness, and not just when they felt sleepy. Within the sample there are a greater number of short haul pilots and as such the frequency of ratings submitted across the entire KSS scale appears to be concomitantly higher. Both short haul and long-haul pilots furthermore were submitting ratings across the 24-hour period, although with a reduced submission rate during the early morning hours (00.00-06.00).

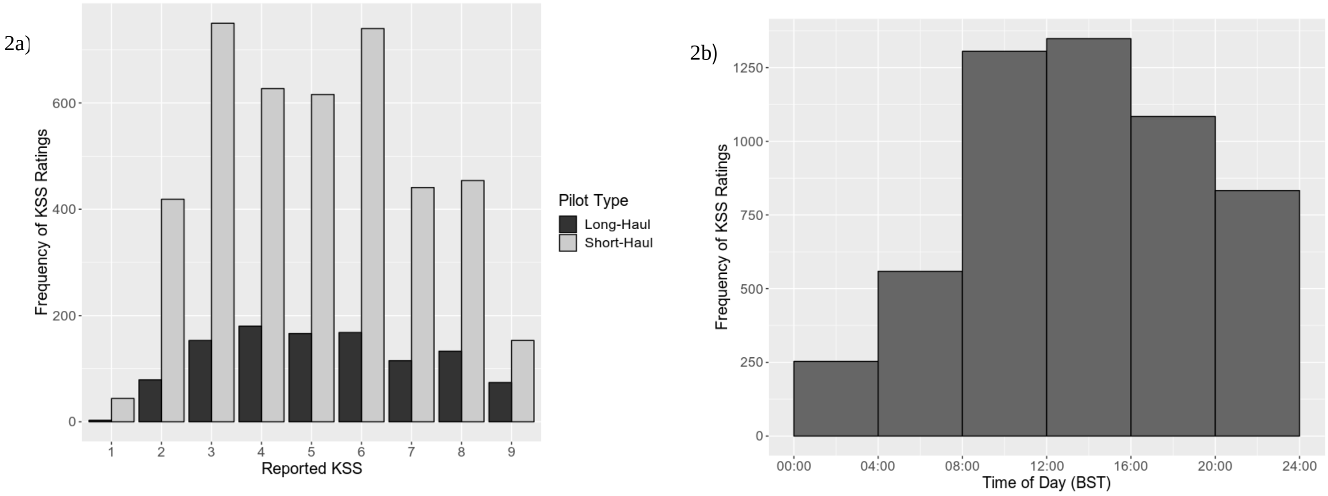

Figure 2a) Overall frequency of KSS ratings during work and non-work hours for long and short haul pilots. 2b) Frequency of KSS ratings plotted in six 4-hour bins across the 24 hour period. Time of day is expressed in British Summer Time (BST).

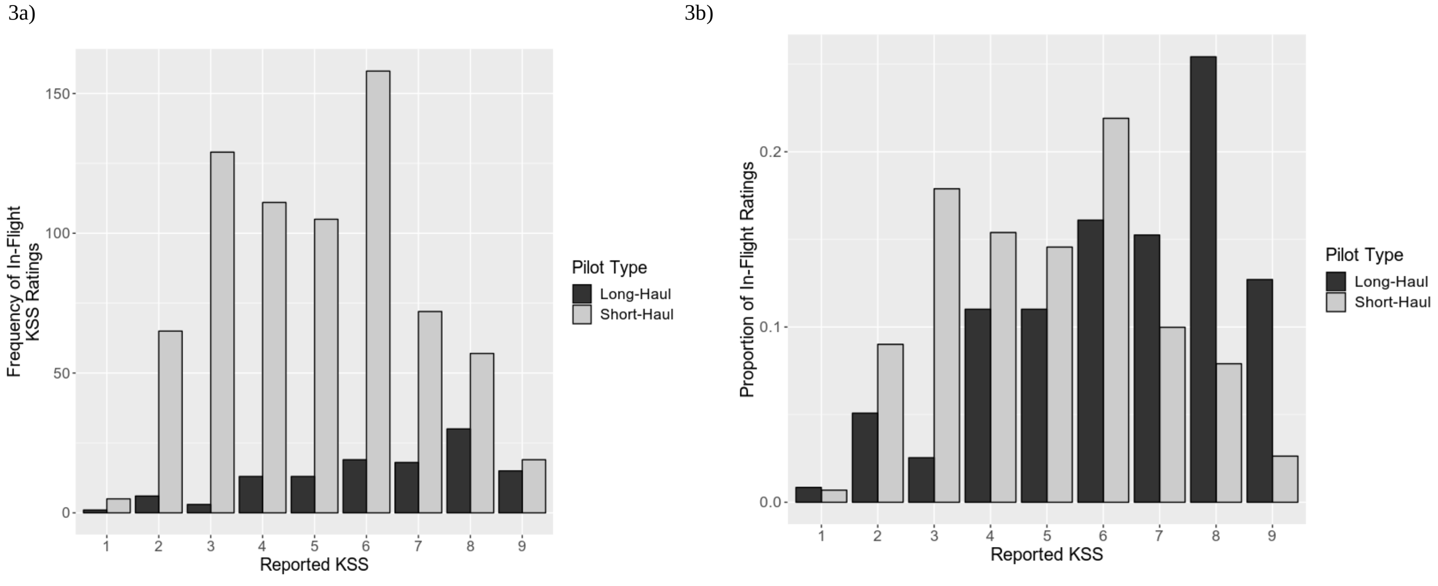

Figure 3a) Frequency of KSS ratings submitted during flying hours by long and short haul pilots. 3b) In-flight KSS ratings expressed as a proportion of all submitted ratings by long and short haul pilots.

As may be seen in 3a) during flying hours, participants submitted 839 KSS ratings, which were distributed across the entire range of the KSS scale for pilots flying both long and short haul operations. There were 211 in-flight KSS ratings (148 from short haul pilots, 63 from long haul pilots) at or above KSS 7 during the one month August period, equating to an overall rate of 3.02 KSS ratings of 7 and above per 100 flying hours. There were 87 KSS 8 ratings (30 from LH pilots, 57 from SH pilots) and 34 KSS 9 ratings (15 from LH, 19 from SH), which equated to an overall rate of 1.7 reports at or above KSS 8 per 100 flying hours. As may be seen in Figure 3b), whilst both long and short haul pilots submitted high KSS scores during flight, long haul pilots submitted a disproportionately high number of KSS 8 and 9 ratings relative to short haul pilots, in line with biomathematical model predictions of the greater proportion of long haul flying hours associated with KSS 7 and above.

3.5 Self-reported involuntary sleep during flight

There were 75 reports of involuntary sleep submitted via the 2-Way App during the one month study period from pilots from nine different companies. Four of these reports came from pilots that had not opted into the study, and hence were excluded from further analysis. The remaining 71 reports came from long haul pilots (n=38) and short haul pilots (n=33) from nine different companies. Forty-two reports related to the participant having involuntarily fallen asleep themselves, 29 related to the other member of flight crew having been reported to involuntarily fallen asleep. Within these, there were nine occasions where there were reports of involuntary sleep for both crew members for the same flight. Figure 3a) shows the absolute number of reports of involuntary sleep broken down by anonymised company and pilot flying role. There were 30 reports from captains (including roles such as training captain) and 41 reports from first officers (including senior first officers). In order to calculate the rate of involuntary sleep report per flight hour, pilots' work schedules are required. Of the 71 reports of involuntary sleep collected, 20 were reported by pilots submitting incomplete roster information, and as such the per-flight hour calculations could not include these reports. The rate of involuntary sleep per flying hour is hence calculated as the number of reports (51) divided by the total flying hours (6986.67), which is 0.0073. Put another way, there were 7.3 reports of involuntary sleep on the flight deck per 1000 flying hours. The rate of involuntary sleep events reported for both flight crew during the same flight was 1.1 reports per 2000 flying hours. Figure 3b. shows the reports of involuntary sleep rate adjusted by flying hours, which reveal substantial variation in involuntary sleep rates between different companies. The difference between absolute and relative figures for company B are likely skewed by a small reporting sample for the company, and therefore more likely to be subject to a degree of random error.

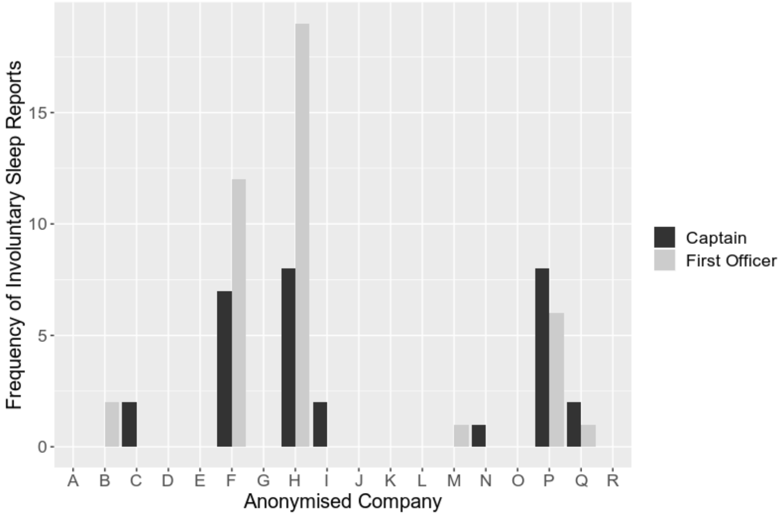

Figure 5a) Absolute frequency of involuntary sleep reports by anonymised company, broken down by piloting role.

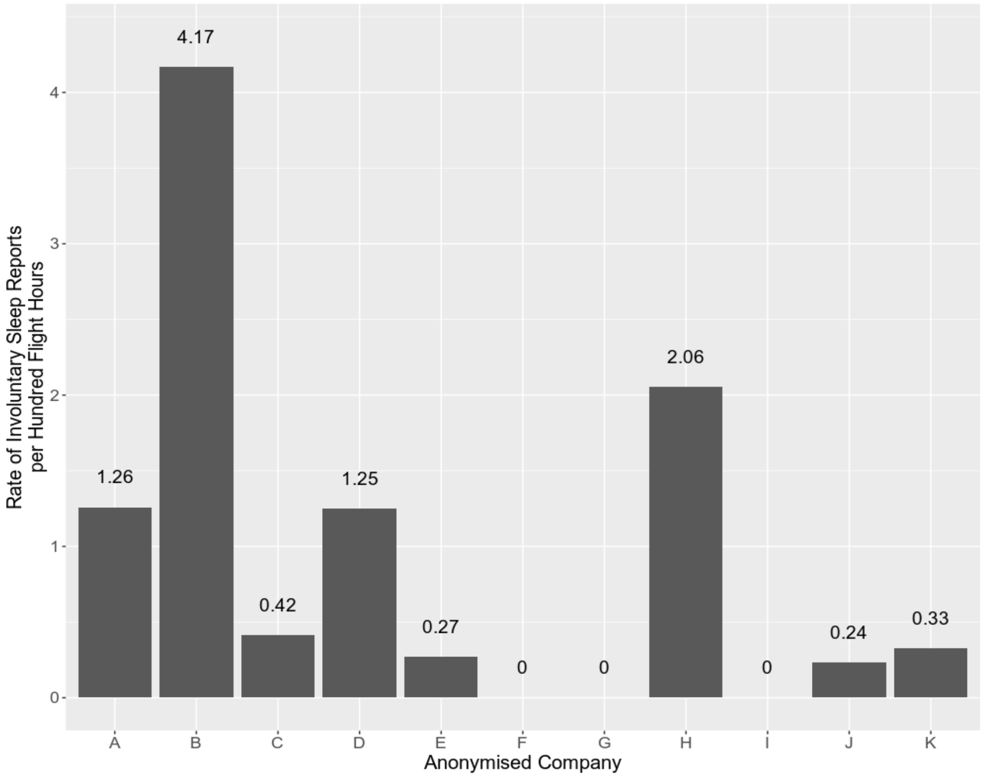

Figure 5b) Rate of involuntary sleep reports per 100 flying hours by anonymised company.

4. Discussion

Against the context of increasing performance-based regulation of the hazard of pilot fatigue in commercial aviation, it is essential to understand the rates of both self-reported and predicted sleepiness levels occurring during flight across the entire aviation industry, using the same scientific measures. Field studies assessing the rates of sleepiness or sleep-related phenomena tied to operational flying hours in aviation are often difficult in practice to conduct with participants from a wide array of different companies and types of operation. The data in the present study provide a benchmark description of predicted fatigue risk rates associated with British airline pilot rosters and actual occurrence rates of high levels of sleepiness during commercial flight.

Predicted sleep opportunities and analysis of duty shift timings prior to duty

A quarter of all flying duty periods were predicted to be preceded by a main sleep opportunity of less than six hours, and up to 43% predicted to be preceded by sleep opportunities of less than seven hours. There is a broad scientific and medical consensus that most healthy adults require between 7-8.5 hours' sleep per night to feel well and maintain full cognitive effectiveness (Kronholm et al., 2009; National Aeronautics and Space Administration, NASA, 1996) Shorter sleep durations of around six hours per night are likely to cause meaningful sleepiness or impaired performance in the average shift worker (Åkerstedt & Wright, 2009; Van Dongen et al., 2003). Against this context, our biomathematical modelling predicted that a substantial proportion of commercial flying duties were preceded by insufficient sleep opportunities. While individuals vary in their need for sleep and their trait vulnerability to the effects of sleep loss (Caldwell et al., 2005; Van Dongen, 2006; Van Dongen & Belenky, 2009) inadequate sleep prior to duty may be a prominent source of schedule-driven fatigue risk exposure for many pilots within UK airline work schedules. As expected, predicted cumulative sleep loss prior to duty as a schedule-driven fatigue risk exposure was almost exclusively seen in short haul operations, with 11.3% of short haul flying duties predicted to be preceded by three consecutive sleep periods of less than seven hours and 2.1% predicted to be preceded by three consecutive sleep periods less than sixhours. Although a relatively much smaller proportion of our sample, consecutive periods of sleep loss present an accumulating fatigue risk for pilots from a safety perspective (Belenky et al., 2003; Van Dongen et al., 2003). Such sleep loss may be of particular concern during operations since chronically sleep-restricted individuals may be less aware of their level of fatigue-related impairment than more acute forms of sleep deprivation (Williamson et al., 2011). Whilst various regulatory principles concerning schedule design emphasize the need for adequate rest in-between duties, and at least an eight hour window of sleep for pilots between duty periods (European Aviation Safety Agency, 2014: ORO.FTL.235), the present findings suggest that in over 40% of actually flown schedules, such ‘out of work' gaps are not predicted to provide eight hour sleep opportunities in terms of their biological plausibility for the average individual.

For both long and short haul pilots, over a quarter of duty start times in this study were shifted three or more hours relative to previous duty start times. As a schedule feature, shifting shift times are likely to interfere with the length and consistency of sleep wake patterns in pilots due to both slow circadian rhythm adaptation to different waking hours, and the difficulty for pilots to adopt consistent coping strategies for abrupt shifts in their sleep patterns within their home lives. Within our sample of 1295 duty periods, duties that started or ended during the known window of circadian low periods furthermore made up over a third of all duties, with 310 (24%) having check in times between 00:00-06:00. The problem with having to get up earlier than usual is that it is very difficult, if not impossible to fall asleep sufficiently early the night before in order to compensate for the early rising time (even when the duty schedule permits), due to lack of adequate homeostasis sleep pressure. Since previous research with pilots and other shift workers has indicated that early duty start timings in particular dramatically restrict the amount of sleep obtained and increase on-duty fatigue levels (Ingre et al., 2008; Roach et al., 2012), this schedule-driven fatigue risk exposure may be a particularly important area for practitioners to target for reduction or provide mitigation measures for within the surrounding duties.

Predicted time of continual wakefulness at last landing

Biomathematical analyses estimated that on average short haul pilots were likely to have been awake for 11.56 hours at last landing, (s.d. = 3.76), compared with long haul pilots with an average 18.31 hours, with a large deviation around the mean (s.d. = 5.6). In terms of safety risks, these findings indicated that pilots operating long haul duties may be particularly at risk of sleepiness and fatigue-related performance decrements towards the end of their duties and on their commute home. Extended periods of wakefulness after about 16-18 hours of wakefulness have a profound impact on alertness levels and performance decline and hence are linked to elevated fatigue-related risks in human operators (National Research Council, 2011; Van Dongen et al., 2003; Williamson et al., 2011). Extensive evidence from both road crash statistics and driving simulator studies further suggest that this elevated fatigue risk exposure at last landing may be important for pilots during their commute home, particularly where duty ends coincide with circadian lows (Horne & Reyner, 1999; Ingre et al., 2006; Reyner & Horne, 1998). In the present study biomathematical estimations of continual hours of wakefulness did not apportion in-flight rest opportunities, an assumption that was estimated to be relevant to 5% of the flying duties analysed in this study.

However, in terms of approximating fatigue-related exposures from work schedules, it was felt that a greater source of inaccuracy would stem from modelling in-flight sleep opportunities where the timing, availability and utility of such opportunities was not known. Future investigations would benefit from the collection of precise in-flight rest data in terms of both possible sleep opportunities and whether such opportunities resulted in sleep across a variety of longer operations. There are rules and principles governing pilots’ hours of work so that they avoid 18 hours of continuous wakefulness during their duties (Civil Aviation Authority, 2016: GM1 CS FTL.1.225(b)(2)). However, our modelling analyses (not taking into account any diversity of sleep-wake strategies prior to duty), still estimated that continual wakefulness associated with pilot work schedules may be a prominent fatigue risk exposure in actually flown schedules, particularly in long haul operations.

KSS Predictions and self-report ratings during Flying Hours

Whilst the average predicted sleepiness scores at key operational points (duty start, last landing and duty end) were overall not indicative of severe KSS sleepiness scores (KSS >7), a substantial proportion of flying hours were associated with predicted and reported sleepiness at levels that may represent a risk to flight safety. For research informing the scale and severity of fatigue hazards experienced by pilots during flight, it is important to underscore what such KSS levels mean. At KSS levels above 7, laboratory research has shown that there is a marked increase in EEG-related sleep intrusions and long eye lid closures, suggesting that at this level, it is difficult, even with high motivation levels, to stave off sleep (Åkerstedt et al., 2014; Anund et al., 2009; Ingre et al., 2006; Reyner & Horne, 1998). KSS levels of 8 and above are of particular concern since this level on the alertness-sleepiness continuum isassociated with markedly increased microsleeping or involuntary sleep intrusion risk, escalating performance decline and increased collision risk in other domains (Åkerstedt et al., 2014). Against this context, the present study findings that pilots submitted 1.7 in-flight KSS ratings at or above KSS 8 per 100 flying hours indicated that high levels of sleepiness in flight occur, and may occur far more routinely than had previously been documented. When compared against the biomathematical model predictions, the present study findings suggest that 10% of flying hours were associated with KSS predictions of 7 or above, within which 3% were associated with predicted fatigue levels of KSS 8 or above. Reported and predicted KSS per-flight hour rates are not directly comparable, since KSS predictions per flying hour can be derived from continuous model predictions, whilst the reported KSS rates of elevated sleepiness levels per flying hour represent discrete occasions where the participants both felt subjectively sleepy *and* submitted a rating. Hence, unless comparable intervals of ratings and predictions are mandated in the methodology, per-flight hour reported rates are likely to be a subset of the predicted KSS per-flight hour rates, because opportunities to complete a rating are not always available, convenient or safe. Notwithstanding this qualification, it is worth noting that both predicted and reported KSS levels of elevated sleepiness (3 hours out of 100 flying hours predicted to be at or above KSS 8 verses a report rate of 1.7 ratings out of 100 flying hours at KSS 8 and above) are of a similar magnitude, and appear high from a safety point of view, given the neurobehavioural deficits evidenced through laboratory work at these levels of sleepiness.

Comparisons of fatigue risks with parallel medical standards for the human component in aviation

As set out in the introduction, within the commercial aviation industry there has not been universal agreement at what rate of occurrence reported or predicted fatigue risks such as these become unacceptable from a broader safety point of view, for a number of reasons. First, there is limited experimental evidence investigating sleep loss, both chronic and acute, and its impact on multi-crew commercial flight performance. Due to advances in the high reliability standards of aircraft automated systems that help control the trajectory of flight for most of the cruise portion, further research is certainly needed to better understand the relationships both between increasing pilot sleepiness levels and unsafe individual pilotperformance, and the resultant impact on the overall safety of flight. Acceptable levels of schedule-driven fatigue risk will furthermore inevitably sit in conflict with commercial productivity and optimisation of crew. As such, the risk appetites of different industry stakeholders regarding acceptability of elevated KSS predictions during flight can, and often do differ within both prescriptive and performance-based regulatory limits (European Aviation Safety Agency, 2014). However, where parallel standards of the risk of incapacitation of flight crew do exist is in the medical incapacitation rate, which is set at a target rate of less than one occurrence per 1,000,000 hours. We found that 10% of flight hours were predicted to be associated with KSS 7 and above, and 3% of flight hours predicted to be associated with KSS 8. This highlights a fundamental disparity between the acceptable probabilities of schedule-driven on-duty fatigue risk, and medical incapacitation risks, despite both risk rates relating to significant in-flight functional impairment of crew. In terms of reported rates, the present study findings of 1.7 KSS ratings of 8 and above per 100 flying hours also appear high against this standard, even though sleepiness may be somewhat more ‘reversible’ by sleep, if this is possible during flight, than many types of medical incapacitation. Indeed, sleep-driven ‘microsleeping’ or involuntary sleep attack events are both the most direct consequence of physiological sleepiness, and also referred to as categories of in-flight medical impairments (Dejohn, Wolbrink, & Larcher, 2004), since the functional impairment to pilot performance is likely to be significant at very high levels of sleepiness (e.g. degradation of visual awareness, attentional lapses and sleep attacks). Within our study there were 71 reports of involuntary sleep during the one month period, and with the data available to calculate a per-flying hour rate, this equated to 7.3 reports of involuntary sleep per 1000 flying hours. Compared with the target medical incapacitation regulation standard of no more than 1 occurrence per 1,000,000 hours, we would argue the magnitude of this approximate 7,000 fold difference is a non-trivial difference in occurrence rate. On four of these occasions involuntary sleep was reported for both flight crew during the same flight, equating to a rate of possible overlapping involuntary sleep events and simultaneous sleepiness in both flight crew of 1.1 reports per 2000 flying hours. The recording of such events are reliant on self-report, and as such it is likely this rate reflects an underestimate of actual occurrences as previous research using sensitive objective recording methods of sleepiness such as EEG in both flight crew has revealed that there may be a number of occasions where pilots involuntarily fall asleep and reawaken without knowledge of the event (Wright & McGown, 2001). Clearly, simultaneous sleepiness within both flight crew is of elevated safety concern since, from a safety systems point of view, as it represents a form of common mode failure (Downer, 2009). This means that the assurance otherwise provided by having independent risks of failure from multiple crew may be compromised where sleepiness risks are similar for pilots flying the same work schedules. Operator fatigue may therefore constitute a more insidious and common source of pilot impairment (Caldwell, 2005; Eriksen et al., 2006; Gander & Signal, 2008; Petrilli et al., 2006), more dangerous than an obvious, complete incapacitation where impaired and potentially unsafe performance goes undetected for an extended period. While a detailed exposition of the medical regulation “1% rule” standard is beyond the scope of this study, it should be mentioned that there is has been some debate as to whether the rule is over- conservative (Mitchell & Evans, 2004). However, the current study findings still highlight a substantial disparity between current acceptable risk standards with respect to pilot functioning and impairment during flight. As some regulatory bodies have pointed out, knowledge of what medical conditions or in-flight impairments are affecting pilots and possibly contributing to a crash or incident '…would be useful in assisting the on-going evolution of the aeromedical regulatory process' (Australian Transport Safety Bureau, 2007: Safety Report B2006/0170, p1-2). As such, it seems important that continued research into pilot fatigue and sleepiness events during actual flying hours are monitored against similar risk standards to other forms of in-flight impairment and incapacitations.

Strengths and limitations

There are several limitations of the present study. First, in order to feasibly sample a larger cohort of pilots and duty patterns, biomathematical model predictions were relied upon to provide the estimates of sleep length and timing, given the pilots' working schedules inputs. Biomathematical models estimate the length and timing of sleep where the biological drives permit based on the duty hours alone, which means that if sleep is predicted to be biologically probable and there is no work duty, it will be apportioned. Hence, in this regard, predictions of sleep length will not take into account other work not detailed in the work schedule, additional travel or out-of-work hassle factors that may extend the individual's hours of wakefulness further. Furthermore, estimations inevitably cannot reflect the diversity of sleeping patterns that pilots may achieve. Whilst additional individual data collection via objective recording instruments such as actigraphy would be desirable, this was not feasible for this initial large-scale exercise, and not considered imperative, since a key aim was to investigate the average likely sleep opportunities that working schedules are likely to provide. Continuous objective measures of sleepiness to complement pilot ratings during flight are desirable, although often not practical for large scale exercises such as the present study. However, analyses of the data showed that participants were, in accordance with instructions, submitting ratings across the KSS scale both during flying hours and non-flying hours, suggesting that the self-report data were not jeopardised by any clear floor or ceiling effects. A further consideration for this study is that it was conducted during the August month period, which is likely to be part of the busy season for many airlines. Hence the work schedules may reflect an increased duty workload for some of the airlines than other seasons within the year. Follow up studies investigating the same variables could indicate if there is substantial seasonal variation across long haul and short haul operations schedule-driven fatigue risks.

The main strength of the present study is the combination of fatigue analysis of flown rosters via biomathematical modelling and collection of self-report data using the same scientific measures across short and long haul pilots from 18 different UK airlines. Such analyses enabled comparisons between the flown work schedules from long and short haul operations and the experience of pilots from different airlines, which together provided insights into the sleepiness risk rates across the entire commercial aviation industry, and not just specific operational routes. Since the large-scale collection of individual sleep habits is often not feasible across an entire workforce, using biomathematical modelling to provide not just on-duty alertness estimations, but also aggregated summaries of predicted sleep and continual wakefulness opportunities prior to flying duties is furthermore useful for providing an overview of sleep-driven exposure risks from a large number of pilot work schedules.

In conclusion, the biomathematical model analyses in the present study indicated that a substantial proportion of flown pilot working schedules are likely to be associated with insufficient sleep opportunities prior to flying duties. High levels of in-flight sleepiness were both reported by pilots and predicted by biomathematical model estimations using the pilots' work schedule times as inputs. Our study provides a benchmark of these rates against the Karolinska Sleepiness Scale. We have suggested that in order to provide context to both predicted fatigue risk rates and self-report rates of high levels of sleepiness, attempts should be made to find common risk denominators, such as per-flight hour rates of occurrence, related to the way that in-flight incapacitations and impairments of pilots are appraised in commercial aviation. Where this comparison is made, the present study findings suggest there is a non-trivial difference of risk tolerance between an existing target medical incapacitation rate for pilots and both the predicted fatigue risk rates associated with pilots' schedules and pilots' reported rates of elevated sleepiness and involuntary sleep during flying hours. Since sleepiness risks during flight are likely to not be independent between flight crew undertaking the same duties, this finding further suggests substantial efforts to reduce fatigue risks during normal flying operations may be required in order to meet the existing target safety standards of the human component.

Acknowledgements

We are grateful for the assistance of the airline pilots who provided their time to participate in this study.

References

Åkerstedt, T., Anund, A., Axelsson, J., & Kecklund, G. (2014). Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. *Journal of Sleep Research*, *23*(3), 240–252. https://doi.org/10.1111/jsr.12158

Åkerstedt, T., Connor, J., & Gray, A. (2008). Predicting road crashes from a mathematical model of alertness regulation — The Sleep / Wake Predictor. *Accident Analysis and Prevention, 40*, 1480–1485. https://doi.org/10.1016/j.aap.2008.03.016

Åkerstedt, T., & Folkard, S. (1996). Predicting duration of sleep from the three process model of regulation of alertness. *Occupational and Environmental Medicine*, *53*(2), 136–141. https://doi.org/10.1136/oem.53.2.136

Åkerstedt, T., & Gillberg, M. (1990). Subjective and Objective Sleepiness in the Active Individual. *International Journal of Neuroscience*, *52*(1–2), 29–37. https://doi.org/10.3109/00207459008994241

Åkerstedt, T., & Wright, K. P. (2009). Sleep Loss and fatigue in Shift Work and Shift Work Disorder. *Sleep Medicine Clinics*, *4*(2), 257–271. https://doi.org/10.1016/j.jsmc.2009.03.001.Sleep

Anund, A., Kecklund, G., Kircher, A., Tapani, A., & Åkerstedt, T. (2009). The Effects of Driving Situation on Sleepiness Indicators after Sleep Loss : A Driving Simulator Study. *Industrial Health*, *47*, 393–401.

Australian Transport Safety Bureau. (2007). Pilot Incapacitation: Analysis of Medical Conditions Affecting Pilots

Involved in Accidents and Incidents. *ATSB TRANSPORT SAFETY REPORT*, (B2006/0170).

Aviation Confidential Incident Reporting Programme (CHIRP). (2017). Air Transport Feedback Issue 121. Retrieved from 121 website: https://www.chirp.co.uk/upload/docs/Air Transport/ATFB 121 - January 2017.pdf

Basner, M., Rao, H., Goel, N., & Dinges, D. F. (2013). Sleep deprivation and neurobehavioral dynamics. *Current Opinion in Neurobiology*, *23*(5), 854–863. https://doi.org/10.1016/j.conb.2013.02.008

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., … Balkin, T. J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *Journal of Sleep Research*, *12*, 1–12.

Borbély, A. A. (1982). A Two Process Model of Sleep Regulation. *Human Neurobiology*, *1*, 195–204. https://doi.org/10.1111/jsr.12371

Borbély, A. A., & Achermann, P. (1992). Concepts and models of sleep regulation: An overview. *Journal of Sleep Research*, *1*(2), 63–79. Retrieved from http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L22219318%5Cnhttp://sfx.library.uu.nl/utrecht?sid=EMBASE&issn=09621105&id=doi:&atitle=Concepts+and+models+of+sleep+regulation%3A+An+overview&stitle=J.+SLEEP+RES.&title=Journal+of+Sle

Bourgeois-Bougrine, S., Carbon, P., Gounelle, C., Mollard, R., & Coblentz, A. (2003). Perceived fatigue for short- and long-haul flights: A survey of 739 airline pilots. *Aviation Space and Environmental Medicine*, *74*(10), 1072–1077.

Cabon, P. H., Coblentz, A., Mollard, R., & Fouillot, J. P. (1993). Human vigilance in railway and long-haul flight operation. *Ergonomics*, *36*(9), 1019–1033. https://doi.org/10.1080/00140139308967974

Caldwell, J. a. (2012). Crew Schedules, Sleep Deprivation, and Aviation Performance. *Current Directions in Psychological Science*, *21*(2), 85–89. https://doi.org/10.1177/0963721411435842

Caldwell, J. A. (2005). Fatigue in aviation. *Travel Medicine and Infectious Disease*, *3*, 85–96. https://doi.org/10.1016/j.tmaid.2004.07.008

Caldwell, J. A. (2012). Crew Schedules, Sleep Deprivation, and Aviation Performance. *Current Directions in Psychological Science*, *21*(2), 85–89. https://doi.org/10.1177/0963721411435842

Caldwell, J. A., Mallis, M. M., Caldwell, J. L., Paul, M. A., Miller, J. C., & Neri, D. F. (2009). Fatigue countermeasures in aviation. *Aviation Space and Environmental Medicine*, Vol. 80, pp. 29–59. https://doi.org/10.3357/ASEM.2435.2009

Caldwell, J. A., Mu, Q., Smith, J. K., Mishory, A., Caldwell, J. L., Peters, G., … George, M. S. (2005). Are individual differences in fatigue vulnerability related to baseline differences in cortical activation? *Behavioral Neuroscience*, *119*(3), 694–707. https://doi.org/10.1037/0735-7044.119.3.694

Cheng, P., & Drake, C. (2016). Occupational Sleep Medicine. *Sleep Medicine Clinics*, *11*(1), 65–79. https://doi.org/10.1016/j.jsmc.2015.10.006

Civil Aviation Authority. (2016). CAP 1265 - EASA Flight Time Limitations: GM1 CS FTL.1.225(b)(2). Retrieved February 15, 2019, from http://publicapps.caa.co.uk/docs/33/CAP 1265 OCT16.pdf

Civil Aviation Authority Safety Regulation Group. (2003). *CAA Paper 2003/04. Wakefulness on the Civil Flight Deck: Evaluation of a Wrist-worn Alertness Device*. 7–13. Retrieved from http://publicapps.caa.co.uk/docs/33/CAPAP2003_14.PDF

Civil Aviation Safety Authority Australia. (2014). *Biomathematical Fatigue Model Guidance Document*. Retrieved from http://wwwtest.casa.gov.au/wcmswr/_assets/main/aoc/fatigue/fatigue_modelling.pdf

ComRes. (2013). ComRes Report prepared on behalf of the British Airline Pilots' Association: BALPA 2013 Membership Survey. Retrieved January 7, 2018, from https://www.comresglobal.com/wp-content/themes/comres/poll/BALPA_Membership_Survey.pdf

Dawson, D., & McCulloch, K. (2005). Managing fatigue: It's about sleep. *Sleep Medicine Reviews*, *9*, 365–380. https://doi.org/10.1016/j.smrv.2005.03.002

Dawson, D., Noy, I. Y., Härmä, M., Åkerstedt, T., & Belenky, G. (2011). Modelling fatigue and the use of fatigue models in work settings. *Accident Analysis and Prevention*, *43*(2), 549–564. https://doi.org/10.1016/j.aap.2009.12.030

Dejohn, C. A., Wolbrink, A. M., & Larcher, J. G. (2004). *FAA Report: In-Flight Medical Incapacitation and Impairment of U . S . Airline Pilots : 1993 to 1998*. Retrieved from https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/media/0416.pdf

Dinges, D. F., & Kribbs, N. B. (1991). Performing while sleepy: Effects of experimentally-induced sleepiness. In *Human Performance and Cognition. Sleep, sleepiness and performance.* (pp. 97–128). Oxford, England: John Wiley & Sons.

Dorrian, J., Darwent, D., Dawson, D., & Roach, G. D. (2012). Predicting pilot's sleep during layovers using their own behaviour or data from colleagues: Implications for biomathematical models. *Accident Analysis and Prevention*, *45*(SUPPL.), 17–21. https://doi.org/10.1016/j.aap.2011.09.019

Downer, J. (2009). *When Failure is an Option : Redundancy, reliability and regulation in complex technical systems*.

Drury, D. A., Ferguson, S. A., & Thomas, M. J. W. (2012). Restricted sleep and negative affective states in commercial pilots during short haul operations. *Accident Analysis and Prevention*, *45*(SUPPL.), 80–84. https://doi.org/10.1016/j.aap.2011.09.031

Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive Consequences of Sleep Deprivation. *Seminars in Neurology*, *25*(1), 117–129. Retrieved from http://faculty.vet.upenn.edu/uep/user_documents/dfd3.pdf

Eriksen, C. A., Åkerstedt, T., & Nilsson, J. P. (2006). Fatigue in trans-atlantic airline operations: Diaries and actigraphy for two- vs. three-pilot crews. *Aviation Space and Environmental Medicine*, *77*(6), 605–612.

European Aviation Safety Agency. (2014a). ORO.FTL.235. Retrieved February 2, 2019, from EU Commission Regulation NO 83/2014 website: https://understandingeasa2016ftl.wordpress.com/easa-ftl/oro-ftl/oro-ftl-235-

rest-periods/

European Aviation Safety Agency, E. (2014b). 1st Workshop on the Approval and Oversight of Fatigue Risk Management (FRM). *The Approval and Oversight of Fatigue Risk Management (FRM)*. Retrieved from https://www.easa.europa.eu/newsroom-and-events/events/1st-workshop-approval-and-oversight-fatigue-risk-management-frm

European Cockpit Association. (2012). ECA fatigue survey barometer. Retrieved January 7, 2018, from https://www.eurocockpit.be/sites/default/files/eca_barometer_on_pilot_fatigue_12_1107_f.pdf

Gander, P. H., Mulrine, H. M., van Den Berg, M. J., Smith, A. A. T., Signal, T. L., Wu, L. J., & Belenky, G. (2014). Pilot fatigue: Relationships with departure and arrival times, flight duration, and direction. *Aviation Space and Environmental Medicine, 85*(8), 833–840. https://doi.org/10.3357/ASEM.3963.2014

Gander, P. H., Signal, T. L., van den Berg, M. J., Mulrine, H. M., Jay, S. M., & Jim Mangie, C. (2013). In-flight sleep, pilot fatigue and Psychomotor Vigilance Task performance on ultra-long range versus long range flights. *Journal of Sleep Research*, *22*(6), 697–706. https://doi.org/10.1111/jsr.12071

Gander, P., & Signal, L. (2008). Who is too old for shift work? Developing better criteria. *Chronobiology International*, *25*(2–3), 199–213. https://doi.org/10.1080/07420520802077556

Goel, N., Basner, M., Rao, H., & Dinges, D. F. (2013). Circadian rhythms, sleep deprivation, and human performance. In *Progress in Molecular Biology and Translational Science* (1st ed., Vol. 119). https://doi.org/10.1016/B978-0-12-396971-2.00007-5

Goode, J. H. (2003). Are pilots at risk of accidents due to fatigue? *Journal of Safety Research*, *34*, 309–313. https://doi.org/10.1016/S0022-4375(03)00033-1

Graeber, R. C., Rosekind, M. R., Connell, L. J., & Dinges, D. (1990). Cockpit napping. *ICAO Journal*, *10,* 6–10.

Horne, J. A., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*(2), 97–110.

Horne, J., & Reyner, L. (1999). Vehicle accidents related to sleep : a review. *Occupational and Environmental Medicine*, *56,* 289–294.

Houston, S., Dawson, K., & Butler, S. (2012). Fatigue reporting among aircrew: Incidence rate and primary causes. *Aviation Space and Environmental Medicine, 83*(8), 800–804. https://doi.org/10.3357/ASEM.3238.2012

Ingre, M., Åkerstedt, T., Björn, P., Anund, A., Kecklund, G., & Pickles, A. (2006). Subjective sleepiness and accident risk avoiding the ecological fallacy. *Journal of Sleep Research, 15,* 142–148.

Ingre, M., Kecklund, G., Åkerstedt, T., Söderström, M., & Kecklund, L. (2008). Sleep length as a function of morning shift-start time in irregular shift schedules for train drivers: Self-rated health and individual differences. *Chronobiology International*, *25*(2–3), 349–358. https://doi.org/10.1080/07420520802110704

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

https://doi.org/10.1371/journal.pone.0108679

International Civil Aviation Organisation. (2006). Annex 1 to the Convention of the International Civil Aviation: Personnel Licensing. In *International Civil Aviation Organization* (pp. 59–73). Retrieved from https://www.theairlinepilots.com/forumarchive/quickref/icao/annex1.pdf

International Civil Aviation Organisation, & ICAO. (2012). *Manual of Civil Aviation Medicine: Doc 8983 AN/895* (3rd ed.).

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006a). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology*, *117*(7), 1574–1581. https://doi.org/10.1016/j.clinph.2006.03.011

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006b). Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology*. https://doi.org/10.1016/j.clinph.2006.03.011

Kandelaars, K., Dorrian, J., Fletcher, A., Roach, G. D., & Dawson, D. (2005). A review of biomathematical fatigue models : Where to from here ? *Sixth International Conference on Fatigue Management in Transportation*, (May), 3–16.

Kato, H., Seguchi, R., Ushijima, T., & Hospital, M. (1993). *NTSB Aircraft Accident Report: AAR-94/04 Uncontrolled Collision With Terrain*. 2–78.

Kecklund, G., & Åkerstedt, T. (1993). Sleepiness in long distance truck driving: An ambulatory EEG study of night driving. In *Ergonomics* (Vol. 36). https://doi.org/10.1080/00140139308967973

Kerkhof, G. ., & Van Dongen, H. P. . (1996). Morning-type and evening-type individuals differ in the phase position of their endogenous circadian oscillator. *Neuroscience Letters*, *218*, 153–156.

Kronholm, E., Sallinen, M., Suutama, T., Sulkava, R., Era, P., & Partonen, T. (2009). Self-reported sleep duration and cognitive functioning in the general population. *Journal of Sleep Research*, *18*, 436–446. https://doi.org/10.1111/j.1365-2869.2009.00765.x

Lyman, G. E., & Orlady, H. W. (1981). NASA Fatigue and Associated Performance Decrements in Air Transport Operations. *NASA Technical Reports Server: 19810014163*, 1–32. https://doi.org/10.1111/add.14038

McCauley, P., Kalachev, L. V., Mollicone, D. J., Banks, S., Dinges, D. F., & Van Dongen, H. P. A. (2013). Dynamic Circadian Modulation in a Biomathematical Model for the Effects of Sleep and Sleep Loss on Waking Neurobehavioral Performance. *Sleep*, *36*(12), 1987–1997. https://doi.org/10.5665/sleep.3246

Mitchell, S. J., & Evans, A. D. (2004). Flight Safety and Medical Incapacitation Risk of Airline Pilots. *Aviation Space and Environmental Medicine*, *75*(3 SEC. I), 260–268.

National Aeronautics and Space Administration, & NASA. (1996). Principles and Guidelines for Duty and Rest Scheduling in Commercial Aviation. *Technical Memorandum*, *110404*, 2–3.

National Research Council. (2011). *The Effects of Commuting on Pilot Fatigue*. https://doi.org/10.17226/13201

National Transportation Safety Board. (2009). *NTSB/AAR-10/01 Accident Report: Loss of Control on Approach Colgan Air Flight 3407*. Retrieved from https://www.ntsb.gov/investigations/reports_aviation.html

National Transportation Safety Board. (2017). *NTSB/AIR-18/01 Incident Report: Taxiway Overflight Air Canada Flight 759*. Retrieved from https://www.ntsb.gov/investigations/AccidentReports/Reports/AIR1801.pdf

National Transportation Safety Board. (2019). Most Wanted List: Reduce Fatigue-Related Accidents.

Petrie, K. J., Powell, D., & Broadbent, E. (2004). Fatigue self-management strategies and reported fatigue in international pilots. *Ergonomics*, *47*, 461–468. https://doi.org/10.1080/0014013031000085653

Petrilli, R. M., Roach, G. D., Dawson, D., & Lamond, N. (2006). The sleep, subjective fatigue, and sustained attention of commercial airline pilots during an international pattern. *Chronobiology International*, *23*(6), 1357–1362. https://doi.org/10.1080/07420520601085925

Phillips, R. O. (2015). A review of definitions of fatigue - And a step towards a whole definition. *Transportation Research Part F: Traffic Psychology and Behaviour*, *29*, 48–56. https://doi.org/10.1016/j.trf.2015.01.003

Pouliquen, Y., Ferrante, O., Jouniaux, G., Cabon, P., Rome, F., Wolff, M., & Mollard, R. (2005). A Human Factors Approach for the Analysis and the Encoding of Aviation Accidents and Incidents : A Validation Study. *2005 International Symposium on Aviation Psychology*, 578–583. Retrieved from https://corescholar.libraries.wright.edu/isap_2005/100

Reis, C., Mestre, C., & Canhão, H. (2013). Prevalence of fatigue in a group of airline pilots. *Aviation Space and Environmental Medicine*, *84*, 828–833. https://doi.org/10.3357/ASEM.3548.2013

Reyner, L., & Horne, J. (1998). Falling asleep whilst driving : are drivers aware of prior sleepiness ? *International Journal of Legal Medicine*, *111*, 120–123.

Roach, G. D., Sargent, C., Darwent, D., & Dawson, D. (2012). Duty periods with early start times restrict the amount of sleep obtained by short-haul airline pilots. *Accident Analysis and Prevention*, *45*(SUPPL.), 22–26. https://doi.org/10.1016/j.aap.2011.09.020

Rosekind, M., Co, E., Gregory, K., & Miller, D. (2000). Crew factors in flight operations XIII: a survey of fatigue factors in corporate/executive aviation operations. *NASA Technical Memorandum*, (September). Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Crew+Factors+in+Flight+Operations+XIII+:+A+Survey+of+Fatigue+Factors+in+Corporate+/+Executive+Aviation+Operations#0

Rosekind, M. R., Co, E. L., Gregory, K. B., & Miller, D. L. (2000). Crew factors in flight operations XIII: a survey of fatigue factors in corporate/executive aviation operationsFIles/BALPA GENERAL/12. PhD1/Fatigue-related Accident reports/SAIB Switzerland Accident AVRO 146-R. *NASA Technical Memorandum*, *209610*(NASA/TM), 3, 26. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Crew+Factors+in+Flight+Operations+XIII+:+A+Survey+of+Fatigue+Factors+in+Corporate+/+Executive+Aviation+Operations#0

Rosekind, M. R., Gander, P. H., Miller, D. L., Gregory, K. B., Smith, R. M., Weldon, K. J., … Lebacqz, J. V.

(1994). Fatigue in operational settings: examples from the aviation environment. *Human Factors*, *36*, 327–338. https://doi.org/10.1177/001872089403600212

Sagaspe, P., Taillard, J., Åkersted, T., Bayon, V., Espié, S., Chaumet, G., … Philip, P. (2008). Extended driving impairs nocturnal driving performances. *PLoS ONE*, *3*(10), 1–6. https://doi.org/10.1371/journal.pone.0003493

Samel, A., Wegmann, H. M., & Vejvoda, M. (1997). Aircrew fatigue in long-haul operations. *Accident Analysis and Prevention*, *29*(4 SPEC. ISS.), 439–452.

Shen, J., Barbera, J., & Shapiro, C. M. (2006). Distinguishing sleepiness and fatigue: Focus on definition and measurement. *Sleep Medicine Reviews*, *10*(1), 63–76. https://doi.org/10.1016/j.smrv.2005.05.004

Swiss Aircraft Accident Investigation Bureau. (2001). *SAIB Final Report No . 1793*. 1–58.

Tunstall-Pedoe, H. (1988). The Concept of Risk. *European Heart Journal*, *9*, 13–15. https://doi.org/10.1111/j.2041-6962.1978.tb01330.x

UK Civil Aviation Authority. (2019). EASA FTL Regulations Combined Document and CAA Guidance to Developing an FTL Scheme. In *EASA Flight Time Limitation Regulations* (pp. 1–69).

Van Dongen, H. P. (2006). Shift work and inter-individual differences in sleep and sleepiness. *Chronobiology International*, *23*(6), 1139–1147. https://doi.org/10.1080/07420520601100971

Van Dongen, H P A, Maislin, G., Mullington, J. M., & Dinges, D. F. (2003). The cumulative cost of additional wakefulness. *Sleep (Rochester)*, *26*, 117–126. https://doi.org/10.1093/sleep/26.2.117

Van Dongen, Hans P A. (2004). Comparison of mathematical model predictions to experimental data of fatigue and performance. *Aviation Space and Environmental Medicine*, *75*.

Van Dongen, Hans P A, & Belenky, G. (2009). Individual Differences in Vulnerability to Sleep Loss in the Work Environment. *Industrial Health*, *47*, 518–526.

Williamson, A., Lombardi, D. A., Folkard, S., Stutts, J., Courtney, T. K., & Connor, J. L. (2011). The link between fatigue and safety. *Accident Analysis and Prevention*, *43*(2), 498–515. https://doi.org/10.1016/j.aap.2009.11.011

Wilson, G. F., Caldwell, J. A., & Russell, C. A. (2007). Performance and Psychophysiological Measures of Fatigue Effects on Aviation Related Tasks of Varying Difficulty. *The International Journal of Aviation Psychology*, *17*(2), 219–247. https://doi.org/10.1080/10508410701328839

Wright, N., & McGown, A. (2001). Vigilance on the civil flight deck: Incidence of sleepiness and sleep during long-haul flights and associated changes in physiological parameters. *Ergonomics*, *44*(1), 82–106. https://doi.org/10.1080/00140130150203893

Zio, E., Fan, M., Zeng, Z., & Kang, R. (2019). Application of reliability technologies in civil aviation: Lessons learnt and perspectives. *Chinese Journal of Aeronautics*, *32*(1), 143–158. https://doi.org/10.1016/j.cja.2018.05.014