A Large-Scale European Union Study of Aircrew Fatigue During Long Night and Disruptive Duties

Mikael Sallinen; Henk van Dijk; Daniel Aeschbach; Anneloes Maij; Torbjörn Åkerstedt

INTRODUCTION: We examined aircrew fatigue during the following flight duty periods (FDPs) mentioned in the European Union (EU) Flight Time Limitations (FTLs): night FDPs longer than 10 h and FDPs typical of disruptive schedules (early starts, late finishes, and nights). An alternative way of classifying night FDPs was also examined to reveal possible subcategories that warrant special attention.

METHODS: A total of 392 aircrew members (96 women) representing 24 airlines participated in the study. Their FDPs were measured by a diary, sleep by the diary and wrist-actigraphy, and fatigue by the Karolinska Sleepiness Scale (KSS) over 14 consecutive days. The KSS ratings given at top of descent (TOD) served as the main outcome.

RESULTS: The probability of high fatigue (KSS $\geq$ 7) at TOD was 0.41 and 0.32 during long ($>$10 h) and short night ($\leq$10 h) FDPs, respectively. The corresponding value was 0.19 for early starts, 0.31 for late finishes, 0.34 for night FDPs, and 0.15 for day FDPs (reference). The main predictors of high fatigue were FDP's encroachment on the window of circadian low (WOCL, 02:00 h–05:59 h) and prior sleep. Within the night category, FDPs fully covering the WOCL showed the highest probability of high fatigue at TOD (0.42).

DISCUSSION: Late finish and night FDPs warrant special attention in fatigue management. Within the night category, the same holds for FDPs that fully cover the WOCL. To manage fatigue, adjustments of the FTLs seem to be a limited strategy and therefore other measures, including maximizing preflight sleep, are needed.

KEYWORDS: pilots, cabin crewmembers, sleepiness, sleep, flight duty period, top of descent.

Sallinen M, van Dijk H, Aeschbach D, Maij A, Åkerstedt T. *A large-scale European Union study of aircrew fatigue during long night and disruptive duties*. Aerosp Med Hum Perform. 2020; 91(8):628–635.

Aircrew fatigue is known to constitute a safety hazard in aviation.[4,10,12] One of the main reasons for fatigue, which can be defined as a biological drive for recuperative rest, is irregular working hours that disturb one's natural sleep-wake cycle. In addition to irregularity, extended durations of flight duty periods (FDPs) often used in civil aviation can amplify aircrew fatigue.

Flight time limitations (FTLs) are the main prescriptive element to protect aircrew from on-duty fatigue. To assess the effectiveness of these limitations in the European Union (EU), the European Union Aviation Safety Agency conducts reviews of them based on operational data. The present study was conducted as part of such a review.[8]

The EU FTL rules identify the following six types of FDPs to be assessed: 1) duties of more than 13 h at the most favorable time of the day; 2) duties of more than 10 h at less favorable times of the day; 3) duties of more than 11 h for crewmembers in an unknown state of acclimatization; 4) duties including a high level of sectors (more than 6); 5) on-call duties such as standby or reserve followed by flight duties; and 6) disruptive schedules.[6] According to the EU FTLs, the term "disruptive schedules" means duty rosters which "disrupt the sleep opportunity during the optimal sleep time window by comprising an FDP or a combination of FDPs which encroach, start or finish during any portion of the day or of the night where a crewmember is acclimatized."[6] In practice, disruptive schedules contain early starts, late finishes, and/or night FDPs.

The present study examined two of the above-mentioned six FDP types: those of more than 10 h at a less favorable time of

From the NLR-Netherlands Aerospace Centre, Amsterdam, The Netherlands.

This manuscript was received for review in November 2019. It was accepted for publication in June 2020.

Address correspondence to: Mikael Sallinen, Finnish Institute of Occupational Health, P.O. Box 40, FI-00033 Työterveyslaitos, Finland; mikael.sallinen@ttl.fi.

DOI: https://doi.org/10.3357/AMHP.5561.2020

the day (i.e., night) and those typical of disruptive schedules. The selection was based on the results of bio-mathematical modeling and an online survey of aircrew. Both these mappings indicated that these two FDP types are the most fatiguing ones out of the six.[8] The focus was on fatigue levels at the top of descent (TOD); that is the point at which the descent to approach altitude is initiated, usually 30 min before landing. The main reason for this selection was that TOD initiates a safety-critical flight phase with high workload.

The first aim of the present study was to assess whether the probability of high fatigue levels at TOD is increased during the FDP types of interest compared to their appropriate reference FDPs, e.g., day FDPs. The focus was on sleepiness type of fatigue, which may be defined as a drive to fall asleep due to, for example, the sleep homeostasis factors and circadian influences.[7,21] The second aim was to identify the main predictors of high fatigue levels at TOD during the FDPs of interest using FDP characteristics (e.g., duration, start time, end time), duration of prior sleep and wake, profession, gender, and other background factors. The third aim was to examine whether an alternative way of classifying current night FDPs would reveal subcategories that are worth considering when grading FDPs according to their effects on aircrew fatigue. This aim was based two observations. First, the night category includes by definition FDPs that either end or start within the window of circadian low (WOCL, 02:00 h–05:59 h) or fully cover the WOCL (i.e., start before and end after the WOCL). Second, the results of the present study suggested that the timing of night FDPs in relation to the WOCL might be a relevant factor for fatigue levels at TOD.

METHODS

Subjects

The study was submitted to the Medical Ethics Review Committee of the Academic Medical Center of the University of Amsterdam. The committee informed the study group that the study did not require an official ethical approval, as the Medical Research Involving Human Subjects act does not apply to our study (reference W17_117.136.). In addition, all data were collected noninvasively and anonymously and all collected flight duty periods would have been undertaken if no study had existed. Each subject signed an informed consent prior to the measurements.

Subjects represented a total of 24 airlines. Six of these airlines had their home base in Eastern Europe, nine in Western Europe, four in Northern Europe, and five in Southern Europe. For the recruitment of airlines, the following two criteria were used to screen for commercial air transport (CAT): the extent to which operators used deviations or derogations from the EU FTL Regulation (operators using such flexibility were excluded) and the types of FDPs used. This resulted in a group of candidate EU CAT operators who were approached and invited to participate. Any other CAT operator could also volunteer to participate. All eligible subjects were made aware of the study and invited to participate by means of a standardized internal email. Most of the airlines translated this email into their own language. In addition, posters informing about the project were hung in the crew rooms.

To participate in this study, an individual volunteer had to be a pilot or cabin crewmember of one of the participating EU CAT operators and operate one or both FDP types of interest. A total of 392 aircrew members [265 airline pilots (24 women) and 127 cabin crewmembers (72 women)] finally participated in the study out of 1634 who originally registered through an online portal which provided them with detailed information on the study. Participating airline pilots and cabin crewmembers were, on average, 40.0 (SD 8.2) and 36.8 (SD 10.1) yr old, respectively. Of the participants, 43.3% of them worked for a network operator, 34.3% for a point-to-point operator, and 21.9% for a cargo operator, and 0.5% for another type of operator.

Procedure

The data were collected between July 2017 and February 2018. Each data collection period, which lasted for 14 consecutive days in maximum, started with 2 d off. On the first day, subjects familiarized themselves with the CrewAlert application (Jeppesen Systems AB, Sweden, part of Boeing Digital Solutions) that was used for data collection. The second day was the first measurement day. During all off-duty periods, besides wearing the actigraph, subjects were asked to rate their sleepiness using the Karolinska Sleepiness Scale (KSS),[2,5] Samn-Perelli (SP) Fatigue Scale,[18] and perform a 5-min version of the Psychomotor Vigilance Task (PVT; pilots only)[14] in the morning, in the afternoon, and in the evening. Subjects were provided with training materials through a dedicated website to familiarize them with the use of the data collection software. They could contact the investigators via telephone or email to get further information.

During flight duty days, subjects were asked to fill in the sleep log, rating their sleepiness on the KSS and the SP using the application soon after waking up and 15 min prior to TOD for each sector. In addition, pilots were asked to perform the PVT soon after awakening and 15 min prior to TOD of the final sector. If an FDP included a long-haul flight, subjects were asked to rate the KSS and the SP also during the cruise phase. With regard to intentional napping during an FDP, the instruction was to press the button on the actigraph in the beginning of the rest period and fill in the sleep log at the end of that period. Napping took always place in the cockpit, as all the measured flights were nonaugmented, which does not permit the pilot to leave the cockpit for a rest break. At the end of each FDP, subjects were asked to fill in information about the completed FDP, their mental effort, and hassle factors. At bedtime, subjects were informed to press the button on the actigraph and fill in the sleep log.

Materials

Because the aim of the study was to examine sleepiness at TOD, only the KSS ratings given at this flight phase will be reported here. The results of the SP and PVT that do not unambiguously represent sleepiness type of fatigue will be reported elsewhere, even though the correlation between the KSS and SP ratings was quite high ($r = 0.868$, $N = 1632$). The collected KSS ratings at TOD were mainly used in a dichotomized form in the analyses, with the ratings 7–9 indicating high fatigue levels (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).

The amount of sleep in the past 24 h prior to TOD and the time elapsed since awakening at TOD were used to describe the sleep/wake patterns prior to and during flight duty days. A dichotomized variable was used to indicate whether subjects took naps during their flights.

The following individual characteristics were collected by a background questionnaire before the field measurements: profession (airline pilot, cabin crew), gender, age, Body Mass Index (BMI), diurnal type,[11] habitual sleep duration per day, and typical commuting time. Based on the sleep log data, the following FDP-related variables were derived: start time, end time, the number of sectors flown, encroachment on the WOCL (yes/no), and the number of time zones crossed either from east to west or from west to east, total FDP duration in 1 wk, and time off prior to an FDP.

Based on FDP start and end times, all FDPs were classified into the following categories according to the current EU FTL regulations:

- early starts: start time between either 05:00 h and 05:59 h (early type of disruptive schedule) or 05:00 h and 06:59 h (late type of disruptive schedule) in the time zone to which the crew is acclimatized;
- late finishes: end time between either 23:00 h and 01:59 h (early type of disruptive schedule) or 00:00 h and 01:59 h (late type of disruptive schedule) in the time zone to which the crew is acclimatized; and
- night FDPs: encroachment on the period between 02:00 h and 04:59 h in the time zone to which the crew is acclimatized.

All FDPs starting at 07:00 h or later and ending before 23:00 h were labeled day FDPs.

Statistical Analysis

The overall probability of high fatigue at TOD (KSS ratings 7–9) in each FDP category of interest, including day FDPs, was calculated in a univariate analysis. The result was the occurrence-probability estimate from all collected FDPs. Furthermore, to get a precise picture of the occurrence of high fatigue at TOD during the FDPs of interest in relation to reference FDPs (day or short night), a subset of the data was extracted from the entire field data to calculate odds ratios (ORs) for high fatigue by means of logistic regression analysis. By this way it was possible to have control over the variation resulting from different numbers of observations per subject in the FDP categories.

For comparisons between long and short night FDPs, a subset of the data was formed by identifying subjects who had one of the two night FDP types measured and those who had both. Subjects with both types of night FDPs measured were randomly assigned to either night FDP group and only the corresponding nights were included. Finally, the first FDP of interest measured was selected from each subject for the analyses, resulting in a subset of the data with a maximum of one observation (short night or long night FDP) per subject.

For comparisons between disruptive and day FDPs, a subset of the data was created by identifying subjects who had the disruptive FDP of interest measured and those who had not. Next, the first FDP of interest measured (either a disruptive FDP or day FDP depending on the FDP group) was selected for the analyses.

To account for profession (pilot or cabin crew), this variable was entered in all analyses as a fixed factor. In addition, all analyses were adjusted for age and gender. To adjust for the FDP-related variables and the other individual-related variables, each of them was entered one at a time as a covariate. Only those variables that reduced a significant *F*-ratio into a nonsignificant ratio are included in the tables.

RESULTS

The occurrence-probability point estimate of high fatigue at TOD during long nights was 0.41 (95% CI:0.34, 0.50) and during short nights 0.32 (95% CI: 0.25, 0.35) in the entire data. Of the long nights, 27% (pilots: 31%, cabin crew: 20%) and of the short nights, 10% (pilots: 11%, cabin crew: 8%) involved on-duty napping.

In the selected subset of the data, the OR for high fatigue at TOD was 1.06 (95% CI: 0.61, 1.85) during long nights when using short night FDPs as a reference. A supplementary analysis based on mean KSS levels at TOD in the same subset of the data showed that the fatigue levels during long and short nights were quite similar, independent of the cutoff used for long duration (9–11 h) (**Fig. 1**). The mean KSS rating at TOD was 5.4 for both long and short nights in all cases.

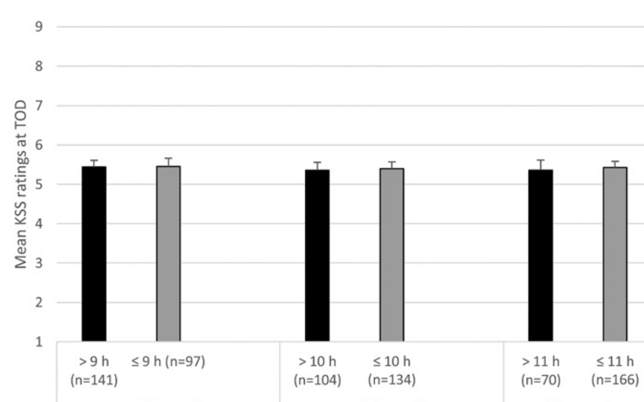

Fig. 1. Mean KSS ratings at TOD for long and short night FDPs. The criterion for the duration of a long night FDP ranged from >9 h to >11 h. The black bars represent long FDPs and the gray bars short FDPs in the selected subsets of the data, including long and short night FDPs with 9 h, 10 h, and 11 h as cutoffs. The vertical lines denote the standard errors.

Fig. 2 shows the occurrence-probability point estimate of high fatigue at TOD for early starts, late finishes, and night FDPs compared to day FDPs in the entire data. This estimate (range 0–1) was 0.18 units greater for night FDPs and 0.16 units greater for late finishes than for day FDPs. The corresponding difference was only 0.04 units for early starts.

On-duty napping was involved in 5% of early starts (pilots: 8%, cabin crew: 1%), 11% of late finishes (pilots: 12%, cabin crew: 8%), and 15% of night FDPs (pilots: 16%, cabin crew: 13%). A supplementary analysis based on the entire data showed that the point estimate of high fatigue for the consecutive FDPs of interest (e.g., 2^{nd}, 3^{rd}, or 4^{th} early start in a row) was 0.07 (95% CI: 0.024, 0.186) for consecutive early starts ($N = 43$), 0.29 (95% CI: 0.13, 0.53) for consecutive late finishes ($N = 17$), and 0.35 (95% CI: 0.26, 0.45) for consecutive nights ($N = 92$). In the selected subset of the data, the OR for high fatigue at TOD was 2.02 (95% CI: 1.09, 3.74) during early starts ($P = 0.026$), 3.77 (95% CI: 2.70, 6.89) during late finishes ($P < 0.001$), and 3.02 (95% CI: 1.81, 5.70) during nights ($P < 0.001$) when using day FDPs as a reference.

First, simple regression analyses were conducted separately for each of the considered independent variables using the subset of the data including both early starts and day FDPs. In this analysis, only FDP start time became a significant predictor of high fatigue at TOD (OR = 0.92, 95% CI: 0.85, 0.99; $P = 0.045$) (**Table I**). After excluding the day FDPs from the data, none of the predictors reached significance.

Table II shows the results of simple regression (Model 1) and multiple regression analyses (Model 2) regarding prediction of high fatigue at TOD in the subset of the data containing both late finishes and day FDPs. In Model 1, longer FDP duration, longer time awake, later FDP end time, and a higher number of time zones crossed from west to east were associated with increased odds of reporting high fatigue, in addition to the late finish FDP type itself. In Model 2 only the significant predictors from Model 2 were entered and it was found that longer FDP duration, later FDP end time, and number of time zones crossed from west to east became significant predictors.

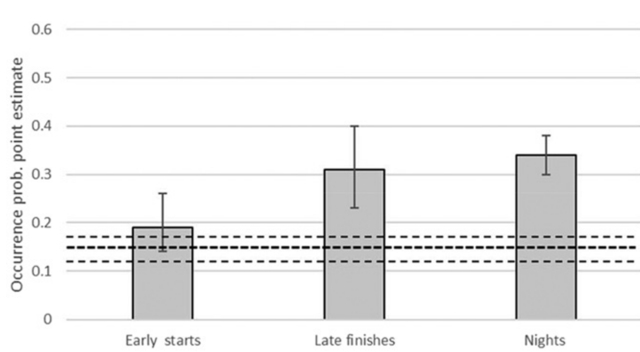

Fig. 2. Point estimates for the probability of high fatigue at TOD during the FDPs of interest and the reference condition (day FDPs). The latter is denoted by the thick dashed horizontal line. The vertical lines indicate 95% CIs of the FDPs of interest. The thin dashed lines denote the 95% CI for day FDPs. The numbers of observations by FDP type are as follows: early starts 163, late finishes 123, nights 496, and days 659.

After excluding the day FDPs from the subset of the data, longer FDP duration (OR = 1.20, 95% CI: 1.04, 1.39, $P = 0.012$), later FDP start time (OR = 0.82, 95% CI: 0.71, 0.94, $P = 0.004$), and later FDP end time (OR = 0.54, 95% CI: 0.31, 0.95, $P = 0.033$) reached significance in Model 1. Model 2, in which all three significant predictors were entered together, yielded an unstable solution because of collinearity. A further analysis was conducted to clarify why later finish times became a protective factor against high fatigue at TOD. The main finding was that the FDPs with earlier end times (23:01 h–00:00 h) were longer in duration (9.7 ± 4 h) than the FDPs with end times between 00:01 h and 01:00 h (7.4 ± 4 h) and those with end times between 01:01 h–02:00 h (9.0 ± 4 h) ($F = 6.1$, $P = 0.003$).

Table III shows the results of simple regression (Model 1) and multiple regression analyses (Model 2) regarding prediction of high fatigue at TOD in the selected subset of the data containing both night and day FDPs. In Model 1, longer FDP duration, WOCL encroachment, lower amount of sleep in the past 24 h, and later FDP start time were associated with increased odds of reporting high fatigue, in addition to the night FDP type itself.

In Model 2, FDP start time was removed even though it was a significant in Model 1. This was done because of collinearity between FDP start time and WOCL encroachment (r = 0.85). Again, WOCL encroachment became the strongest predictor of high fatigue at TOD.

After excluding the day FDPs from the subset of the data, the amount of sleep in the past 24 h (OR = 0.84, 95% CI: 0.75, 0.94, $P = 0.002$) and being a cabin crewmember (vs. airline pilot) (OR = 1.89, 95% CI: 1.07, 3.36, $P = 0.029$) were the significant predictors of high fatigue at TOD in simple logistic regression analyses. When both these predictors were entered simultaneously into a regression analysis, the amount of sleep in the past 24 h still became a significant predictor (OR = 0.79, 95% CI: 0.67, 0.93, $P = 0.004$) but being a cabin crewmember (vs. airline pilot) only approached the level of significance (OR = 2.33, 95% CI: 0.99, 5.44, $P = 0.051$). A supplementary simple regression analysis showed that the number of sectors was the only FDP characteristic to predict high fatigue at TOD among day FDPs (OR = 1.67, 95% CI: 1.03, 3.36, $P = 0.039$).

There were two reasons to further study the night FDPs. First this category included three types of FDPs: those with start times within the WOCL (resembling early starts), those with end times within the WOCL (resembling late finishes), and those with start times before and end times after the WOCL (covering the whole night). Second, the results reported above showed that the probability of high fatigue at TOD was almost the same for late finishes (0.31) and night FDPs (0.34), but clearly lower for early starts (0.19), approaching the level of day FDPs (0.15). These observations suggested that these three subcategories of night FDPs differ in the probability of high fatigue at TOD.

A supplementary analysis was conducted to study the overall role of lateness of FDP end time. For this purpose, the

Table I. Logistic Regression Analysis of Variables Predicting the Odds Ratio (OR) of High Fatigue at TOD.

	MODEL 1* UNADJUSTED OR (CI)	*P*
FDP duration (h)	1.11 (0.99; 1.23)	0.065
No. of sectors	1.04 (0.78; 1.37)	0.803
Time awake prior to TOD (h)	1.05 (1.00; 1.11)	0.051
FDP end time (h)	0.98 (0.92; 1.04)	0.529
Sleep in the past 24 h (h)	0.90 (0.78; 1.02)	0.103
FDP start time (h)	0.92 (0.85; 0.99)	0.045
Profession (cockpit vs. cabin crew, 0/1)	1.36 (0.74; 2.50)	0.321
Gender (man vs. woman, 0/1)	0.96 (0.49; 1.88)	0.957
Age (years)	0.98 (0.95; 1.01)	0.212
Napping during FDP (yes vs. no, 1/0)	0.99 (0.95; 1.04)	0.788
No. of time zones crossed from West to East	1.17 (1.00; 1.34)	0.051
No. of time zones crossed from East to West	0.94 (0.80; 1.11)	0.474
FDP type (early start vs. day 1/0)	1.77 (0.98; 3.20)	0.059

The analysis is based on the selected subset of the data with early starts and day FDPs. $N = 299$.
FDP: flight duty period; TOD: top of descent; OR: odds ratios; CI: confidence interval.
* Simple logistic regression.

selected subset of the data including late finishes and night FDPs was extracted. To avoid mixing night FDPs with and without an opportunity for prior night sleep, only FDPs with start time before midnight were selected. The analysis showed that later end time was associated with a higher probability of high fatigue at TOD ($P = 0.003$) (**Fig. 3**). This suggested that the subcategory of night FDPs with end time after the WOCL shows an exceptionally high probability of high fatigue at TOD.

Fig. 4 shows the occurrence-probability point estimate of high fatigue at TOD for the three subcategories of night FDPs compared to day FDPs in the entire data. The point estimate was highest for the "finish after the WOCL" subcategory and lowest for the "start inside the WOCL" subcategory, the difference being 0.19 units between these two. All these three FDP types showed higher point estimates than day FDPs.

A supplementary analysis showed that the sleep/wake ratio at TOD was lowest for the "finish after the WOCL" FDPs (3.91 ± 2.91 h/17.80 ± 6.81 h), followed by the "finish inside the WOCL" FDPs (4.95 ± 2.52 h/16.75 ± 6.73 h) and was highest for the "start inside the WOCL" FDPs (5.64 ± 2.37 h/11.13 ± 4.89 h). FDP duration was longest for the "finish after the WOCL" FDPs (10.10 ± 2.50 h), followed by the "finish inside the WOCL" FDPs (8.30 ± 2.71 h), and was shortest for the "start inside the WOCL" FDPs (6.49 ± 2.92 h).

In the selected subset of the data including the subcategories of night FDPs and day FDPs, the OR for high fatigue at TOD was 4.16 (95% CI: 1.63–10.22) in the "start inside the WOCL" subcategory ($P < 0.001$), 4.16 (95% CI: 2.00–8.65) in the "finish inside the WOCL" subcategory ($P < 0.001$), and 8.04 (95% CI: 3.58–18.0) in the "finish after the WOCL" subcategory ($P < 0.001$) when using day FDPs as reference.

DISCUSSION

Our field study shows that the probability of rating high fatigue at TOD is increased during night and late finish FDPs as compared with day FDPs in flying personnel. Less pronounced results of increased fatigue at TOD was found for early start FDPs. A similar pattern was found for on-duty napping, with the highest probability occurring during night FDPs and the lowest during early starts. A detailed analysis of night FDPs suggests that long FDP duration is not a significant predictor of high fatigue at TOD during night flights. Finally, the current definition of the night FDP, encroachment on the period between 02:00 h and 04:59 h, seems to encompass a very heterogeneous group of FDPs in terms of high fatigue at TOD. Preliminary findings suggest that high fatigue at TOD is most likely to occur during night FDPs with end time after the WOCL. Of all studied FDP characteristics, WOCL encroachment proved the most powerful to predict high fatigue at TOD.

Table II. Logistic Regression Analysis of Variables Predicting the Odds Ratio of High Fatigue at TOD.

	MODEL 1* UNADJUSTED OR (CI)	*P*	MODEL 2† OR (CI)	*P*
FDP duration (h)	1.23 (1.11;1.37)	0.000	1.13 (1.01;1.27)	0.038
No. of sectors	1.17 (0.89;1.54)	0.261		
Time awake prior to TOD (h)	1.09 (1.04;1.14)	0.000	1.02 (0.97;1.06)	0.499
Sleep in the past 24 h (h)	0.97 (0.83;103)	0.168		
FDP start time (h)	1.03 (0.96;1.10)	0.399		
FDP end time (h)	1.17 (1.08;1.28)	0.000	1.19 (1.06;1.23)	0.003
Profession (cockpit vs. cabin crew, 0/1)	1.45 (0.82;2.58)	0.201		
Gender (man vs. woman, 0/1)	0.90 (0.48;1.69)	0.744		
Age (years)	0.90 (0.48;1.69)	0.744		
Napping during FDP (yes vs. no, 1/0)	0.98 (0.94;1.03)	0.385		
No. of time zones crossed from West to East	1.16 (1.05;1.30)	0.005	1.07 (0.95;1.21)	0.035
No. of time zones crossed from East to West	0.88 (0.74;1.03)	0.116		
FDP type (late finish vs. day, 1/0)	3.21 (1.81;5.68)	0.000	#	

The analysis is based on the selected subset of the data with late finishes and day FDPs. $N = 314$.
FDP: flight duty period; TOD: top of descent; OR: odds ratios; CI: confidence interval.
* Simple logistic regression; †multiple logistic regression; #removed variable.

Table III. Logistic Regression Analysis of Variables Predicting the Odds Ratio (OR) of High Fatigue at TOD.

	MODEL 1* UNADJUSTED OR (CI)	*P*	MODEL 2† OR (CI)	*P*
FDP duration (h)	1.08 (1.01;1.16)	0.037	0.97 (0.88;1.07)	0.561
WOCL encroachment (yes vs. no, 1/0)	3.00 (1.69;5.33)	0.000	3.20 (1.11;9.19)	0.047
No. of sectors	0.91 (0.73;1.14)	0.428		
Time awake prior TOD (h)	1.03 (0.99;1.06)	0.090		
Sleep in the past 24 h (h)	0.84 (0.76;0.91)	0.000	0.86 (0.77;0.96)	0.009
FDP start time (h)	1.06 (1.02;1.10)	0.003	#	
Profession (cockpit vs. cabin crew, 0/1)	1.59 (0.98;2.57)	0.059		
Gender (man vs. woman, 0/1)	0.78 (0.47;1.32)	0.356		
Age (years)	1.00 (0.97;1.02)	0.878		
On-duty napping (yes vs. no, 1/0)	1.00 (0.98;1.03)	0.877		
No. of time zones crossed WE	1.04 (0.95;1.14)	0.426		
No. of time zones crossed EW	1.06 (0.94;1.20)	0.336		
FDP type (night vs. day, 1/0)	3.03 (1.72;5.34)	0.000	#	

The analysis is based on the selected subset of the data with night and day FDPs. $N = 361–335$ depending on variable.
FDP: flight duty period; WOCL: window of circadian low; TOD: top of descent; OR: odds ratios; CI: confidence interval.
* Simple logistic regression; †multiple logistic regression; #removed variable.

The results regarding high fatigue at TOD during night duties are well in line with previous results reported from various industries including aviation.[2,16] A somewhat surprising result was the high probability of fatigue during late finishes, which was similar to the corresponding probability during night FPDs. Usually, fatigue levels measured during evening duties are even lower than those measured during early morning duties.[2] This discrepancy between the present and previous studies can probably be explained by the fact that typically evening duties end at around 22:00 h, whereas in the present study these duties, called late finishes, ended a few hours later, between 23:00 h and 01:59 h. The results, however, are consistent with a study in short-haul pilots where increased fatigue levels were found at the end of late-finishing FDPs.[19] This finding was in part explained by long wake time in connection with these FPDs. Moreover, in most adults, fatigue is known to increase between 23:00 h and 01:59 h due to the start of a downswing of circadian-regulated alertness.[3] A reason for the relatively moderate levels of fatigue during the early start FDPs probably lies in the fact that TOD typically occurs at quite favorable times of the day during these FDPs, i.e., during the upswing of circadian wake promotion. In addition, the time since awakening is still relatively short in this case. Moreover, when comparing early start FDPs with day FDPs, it is worth taking into account that the latter may also restrict prior sleep especially when FDPs start between 07:00 h and 08:00 h. In these cases, rise time is often quite early due to the time needed for morning routine and commuting.

Our results provided no evidence that the odds of reporting high fatigue at TOD would increase with having more than one disruptive FDP in a row. The data did not, however, permit analyzing cumulative fatigue over sequences longer than two consecutive FDPs. There might be a regulatory reason for this limitation. The current prescriptive rules in the EU require airlines to extend the recovery rest period if a crewmember has had four or more disruptive schedules (CS.FTL.1.235 Rest Periods). This rule may limit the use of such scheduling solutions.

There are two interesting aspects to fatigue at TOD during night FDPs: FDP duration and timing in relation to the WOCL. Our result that FDP duration was not a strong determinant of fatigue during night FDPs is somewhat surprising, as long duty

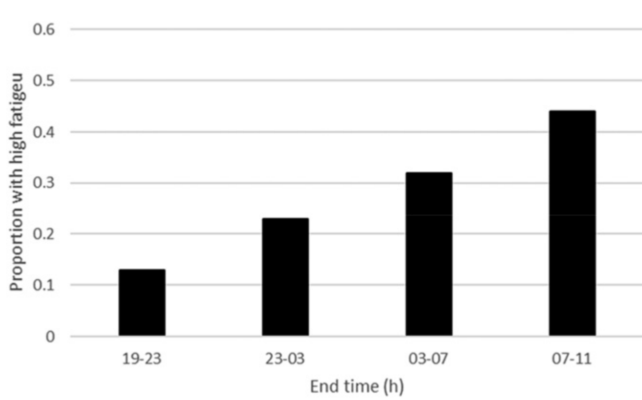

Fig. 3. Proportion of high fatigue levels at TOD plotted against FDP end time in the subset of the data with late finishes and night FDPs. All FDPs start before midnight in this subset of the data.

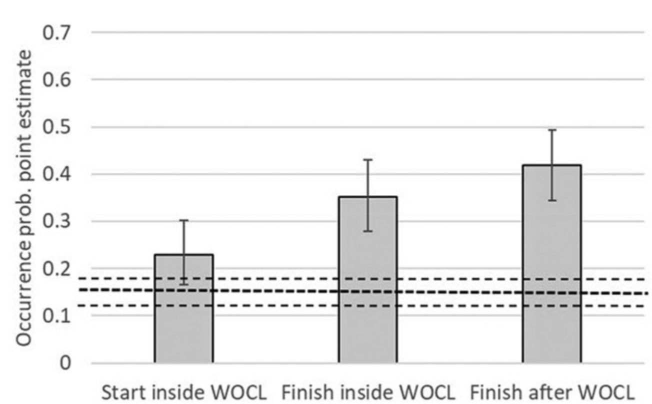

Fig. 4. Point estimates for the probability of high fatigue at TOD during the subcategories of night FDPs and the reference condition (day FDPs). The latter is denoted by the thick dashed horizontal line. The vertical lines indicate 95% CIs of the FDPs of interest. The thin dashed lines denote the 95% CI for day FDPs. The numbers of observations by FDP type are as follows: FDPs with start inside the WOCL 157, FDPs with end inside the WOCL 162, FDPs with finish after the WOCL 177, and days 659.

hours have been found to contribute to fatigue and accident/injury risk.[13,20] The result might be explained by the fact that fatigue during night duties is mainly due to unfavorable time of the day (circadian factor) and an insufficient amount of sleep in relation to the time spent awake (homeostatic factor). In the present study, the effects of these determinants possibly interacted with those of FDP duration. In addition, night FDPs were seldom of short duration (e.g., 1.5% of the night FDPs were 4 h or shorter), which restricts the possibility of studying the role of this FDP characteristic. Also, the length of time subjects were fatigued was not measured, which may also underestimate the role of FDP duration in the present study. Thus, our results should not be used to conclude that FDP duration is an unimportant factor in fatigue mitigation through prescriptive rules. However, our results do suggest that both night FDPs longer than 10 h and shorter than or equal to 10 h should receive equal attention when trying to mitigate fatigue at TOD.

The results regarding the studied subcategories of night FDPs suggest that the timing of a night FDP in relation to WOCL should be taken into account in fatigue management. According to our results, night FDPs with end time after the WOCL deserve the most attention. There are at least two explanations for the result of the high probability of fatigue at TOD during these night FDPs: a low sleep-wake ratio when arriving at TOD and exceptionally long FDP duration.

Our results of on-duty napping indicate that this measure of fatigue mitigation is quite frequently used even during nonaugmented flights during which pilots have no possibility to sleep in a separate rest facility outside the cockpit. This result is in line with a previous study on airline pilots.[17] As all the flights of the present study were nonaugmented flights, on-duty napping may have been used as a countermeasure for unexpected fatigue under the controlled rest procedure among pilots. On the other hand, the high proportion of especially long night FDPs with on-duty napping questions the unexpected nature of fatigue, which in turn gives cause to consider if napping on the flight deck could be used in a planned manner to mitigate fatigue during nonaugmented night flights.

Our results suggest that WOCL encroachment is the FDP characteristic that most strongly predicts high fatigue at TOD. WOCL encroachment approximately tripled the odds of reporting high fatigue at TOD, while for the other studied FDP characteristics, the corresponding odds remained much lower. The significance of WOCL encroachment is in accordance with numerous shiftwork studies showing a strong downswing of circadian-regulated alertness at night.[2,16] In all, this implies that as long as FDPs encroach on the WOCL, it is difficult to effectively manage fatigue just by means of prescriptive rules and, thus, effective fatigue risk management strategies are also needed.

In addition, the present analysis identified longer sleep in the 24 h prior to a night FDP as a protective factor against high fatigue at TOD. This result is in line with a recent study in which airline pilots' recurrent fatigue during nighttime flights was found to be associated with shorter prior sleep.[15] Both these results suggest that the role of sleep prior to night FDPs should be emphasized in fatigue mitigation, in addition to on-duty fatigue mitigation strategies such as in-flight rest.

Apart from WOCL encroachment, the studied FDP characteristics predicted high fatigue at TOD within each FDP category only weakly. Within the early start category, none of the studied FDP characteristics predicted high fatigue at TOD after excluding day FDPs from the analyses. Within the late finish category, the overall picture was mixed. Especially our result of the positive association between later finish time and high fatigue at TOD warrants further research.

In the EU FTL regulations, the number of sectors flown is one of the main FDP characteristics, besides FDP start time and duration and prior rest period, that is used to mitigate fatigue. For example, the maximum FDP duration depends on the number of sectors flown: the more sectors flown, the shorter the maximum FDP duration is allowed to be. In the present study, the number of sectors was not a significant predictor of high fatigue at TOD in any FDP category of interest. This result is in conflict with the results of previous studies.[1,9,13] However, these studies focused on day FDPs, whereas the present study focused on disruptive FDPs. This difference probably explains the conflict, because in the present study, the number of sectors was also a significant predictor of high fatigue at TOD when only day FDPs were included in the analysis.

The difference between airline pilots and cabin crewmembers in fatigue at TOD during night FDPs is of interest. In fact, occupational group was the only individual-related characteristic to significantly predict high fatigue at TOD. One explanation for the difference between the two groups of aircrew lies in the workload before and during TOD. At this flight phase, cabin crew, who reported high fatigue at TOD more frequently than pilots, are typically sitting after a potentially hectic work period, whereas pilots are in a high workload phase of flight after the quieter cruise phase. On the other hand, the differences in workload cannot be the only explanation, since profession predicted high fatigue at TOD neither during early starts nor during late finishes. Thus, further studies are needed to confirm this result and explain it comprehensively, also taking into account differences in nap break opportunities between the occupational groups.

The main strength of the present study is that the results are based on a large dataset collected from the crews of 24 airlines around Europe. Also, the results are based on field data collected under naturalistic working conditions. Thus they reflect the current situation in Europe more reliably than field studies conducted within one or two airlines, or a fatigue survey among flying personnel. The latter does not permit one to reliably assess how sleepy an aircrew really has been at a certain phase of the flight. Nor does it permit a reliable determination of the FDP characteristics within an FDP category that predict high fatigue at a certain phase of a flight. On the other hand, the number of subjects per airline varied greatly in the present study and the sample was not based on randomization. Both these factors limit the representativeness of the study.

Our data combined different types of air transport and operators, which may mask some features of on-duty fatigue specific to only one type of air transport or operator, such as cargo. This is a topic to address in future studies.

Finally, the results should be interpreted bearing in mind that the data did not cover all possible FDPs one can schedule under the current EU FTL rules regarding long night and disruptive FDPs. Thus fatigue levels at TOD for the most extreme (but possible) scheduling solutions remain unknown.

To conclude, FDPs typical of disruptive schedules are associated with a high probability of fatigue at TOD compared to day FDPs, and this phenomenon is especially pronounced during late finish and night FDPs. In addition, night FDPs with end time after the WOCL seem to constitute a subcategory for which individual and organizational strategies to mitigate fatigue are needed. In all, the potential of improving aircrew fatigue mitigation through adjustments of the current EU FTL rules seems rather limited as long as disruptive FDPs are used. Instead, improvements in sleep opportunities prior to night FDPs and on-duty fatigue mitigation strategies are more promising approaches for this purpose.

ACKNOWLEDGMENTS

The authors thank the aircrews who participated in the study and the European Union Aviation Safety Agency for their support throughout the work.

Financial Disclosure Statement: This study was funded by European Commission (contract no. MOVE/B3/SER/2016-360/SI2.747687, "Effectiveness of the provisions concerning flight and duty time limitations and rest requirements contained in Annexes II and III of Commission Regulation No. 965/2012") and NordForsk, Nordic Programme on Health and Welfare (grant no. 74809). The authors have no competing interests to declare.

Authors and affiliations: Mikael Sallinen, Ph.D., Adjunct Professor, Finnish Institute of Occupational Health, Helsinki, Finland; Henk van Dijk, Ph.D., and Anneloes Maij, M.Sc., Ph.D., NLR-Netherlands Aerospace Centre, Amsterdam, The Netherlands; Daniel Aeschbach, Ph.D., Professor, Department of Sleep and Human Factors Research, Institute of Aerospace Medicine, German Aerospace Center, Cologne, Germany; and Torbjörn Åkerstedt, Ph.D., Professor, Stress Research Institute, Stockholm University, Stockholm, Sweden.

REFERENCES

1. Aeschbach D, Vejvoda M, Mendolia F, Tritschler K. Additive effects of the number of completed flights and time awake on fatigue in short-haul airline pilots. Sleep. 2017; 40(Suppl_1):A58.
2. Åkerstedt T, Anund A, Axelsson J, Kecklund G. Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. J Sleep Res. 2014; 23(3):242–254.
3. Åkerstedt T, Folkard S. Validation of the S and C components of the three-process model of alertness regulation. Sleep. 1995; 18(1):1–6.
4. Avers K, Johnson WB. A review of Federal Aviation Administration fatigue research: transitioning scientific results to the aviation industry. Aviat Psychol Appl Hum Factors. 2011; 1(2):87–98.
5. Baulk SD, Reyner LA, Horne JA. Driver sleepiness— evaluation of reaction time measurement as a secondary task. Sleep. 2001; 24(6):695–698.
6. Commission Regulation (EU). No. 83/2014. [Accessed 7 Oct. 2019]. Available from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2014:028:0017:0029:EN:PDF.
7. Dement WC, Carskadon MA. Current perspectives on daytime sleepiness: the issues. Sleep. 1982; 5(Suppl_2):S56–S66.
8. EASA report on Effectiveness of Flight Time Limitation (FTL). 2019. [Accessed 7 Oct. 2019]. Available from https://www.easa.europa.eu/sites/default/files/dfu/Report%20on%20effectiveness%20of%20FTL_final.pdf.
9. Goffeng EM, Wagstaff A, Nordby K-C, Meland A, Goffeng LO, et al. Risk of fatigue among airline crew during 4 consecutive days of flight duty. Aerosp Med Hum Perform. 2019; 90(5):466–474.
10. Hartzler BM. Fatigue on the flight deck: the consequences of sleep loss and the benefits of napping. Accid Anal Prev. 2014; 62:309–318.
11. Horne JA, Östberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol. 1976; 4(2):97–110.
12. International Air Transport Association, International Civil Aviation Organisation, International Federation of Airline Pilots Associations. Fatigue Risk Management Systems Implementation Guide for Operators, 2011. [Accessed 7 Oct. 2019]. Available from https://www.icao.int/safety/fatiguemanagement/FRMS%20Tools/FRMS%20Implementation%20Guide%20for%20Operators%20July%202011.pdf.
13. Powell DMC, Spencer MB, Holland D, Broadbent E, Petrie KJ. Pilot fatigue in short-haul operations: effects of number of sectors, duty length, and time of day. Aviat Space Environ Med. 2007; 78(7):698–701.
14. Roach GD, Dawson D, Lamond L. Can a shorter psycho-motor vigilance task be used as a reasonable substitute for the ten-minute psychomotor vigilance task? Chronobiol Int. 2006; 23(6):1379–1387.
15. Sallinen M, Åkerstedt T, Härmä M, Henelius A, Ketola K, et al. Recurrent on-duty sleepiness and alertness management strategies in long-haul airline pilots. Aerosp Med Hum Perform. 2018; 89(7):601–608.
16. Sallinen M, Hublin C. Fatigue-inducing factors in transportation operations. In: Popkin S, editor. Reviews of Human Factors and Ergonomics: Worker Fatigue and Transportation Safety, vol. 10. Thousand Oaks (CA, USA): Sage Publications; 2015:138–173.
17. Sallinen M, Sihvola M, Puttonen S, Ketola K, Tuori A, et al. Sleep, alertness and alertness management among commercial airline pilots on short-haul and long-haul flights. Accid Anal Prev. 2017; 98:320–329.
18. Samn SW, Perelli LP. Estimating aircrew fatigue: a technique with application to airlift operations. Brooks AFB (TX, USA): USAF School of Aerospace Medicine; 1982. Technical Report SAM-TR-82-21.
19. Vejvoda M, Elmenhorst EM, Pennig S, Plath G, Maass H, et al. Significance of time awake for predicting pilots´ fatigue on short-haul flights: implications for flight duty time regulations. J Sleep Res. 2014; 23(5):564–567.
20. Wagstaff AS, Sigstad Lie JA. Shift and night work and long working hours: a systematic review of safety implications. Scand J Work Environ Health. 2011; 37(3):173–185.
21. Williamson A, Lombardi DA, Folkard S, Stutts J, Courtney TK, Connor JL. The link between fatigue and safety. Accid Anal Prev. 2011; 43(2):498–515.