Layover Sleep Prediction for Cockpit Crews During Transmeridian Flight Patterns

Katie J. Kandelaars, Adam Fletcher, Guy E. Eitzen, Greg D. Roach, and Drew Dawson

SUMMARY

Current models of fatigue and alertness use a combination of biological (circadian) and homeostatic factors to predict sleep and wake. Such models do not include social factors in their calculations. The aim of our analysis was to compare the relative contributions of social and biological factors in models designed to predict the total sleep time (TST) during layover periods between transmeridian flights.

The study actigraphically collected sleep information from 86 cockpit crew (mean age 46.7 yr, SD 4.3 yr) during round-trip patterns from Australia to Los Angeles (n = 15), Europe (n = 42), New York (n = 10), and Hong Kong (n = 19). Linear regression models were constructed to predict TST using data from airline schedules. This schedule information included layover length, flight duration, the number of night hours at the destination (social hours), the number of night hours in Australian Eastern Standard time (biological hours), and time zone displacement. These models were then validated using independent data.

Analysis indicated that the social factors account for more variance than biological factors in their calculation of social night hours account for more variance than biological night hours (r = 0.8 vs. 0.7). Additionally, the layover length achieved a correlation coefficient of 0.9. These results were strengthened when the model parameters were applied to the cross-validation dataset. Social night hours significantly influence sleep during international layovers and may be a better predictor than biological night hours. More research must be carried out to determine the validity of these findings in a larger, randomly collected flight sample.

Keywords: sleep prediction, transmeridian flight, social impact, biological impact.
and hence fatigue and/or alertness levels are not easy to estimate. A number of mathematical models have been formulated to predict fatigue and alertness in operational settings, with summaries and a comparison of these models recently published in *Aviation, Space, and Environmental Medicine* (Vol. 75, No. 3, Suppl.). These models usually require actual sleep and wake data to predict the level of fatigue associated with different schedules. However, this information is not known prospectively; therefore, some models estimate sleep and wake times based on physiological parameters such as circadian phase and the homeostatic drive for sleep (5). Currently, no approach described in the literature accounts for geophysical factors such as the impact of local night onset while flying transmeridian flight patterns. From a fatigue and alertness perspective, one recent study has described a dose-response effect between lapses in behavioral alertness and the cumulative duration of wakefulness in excess of 15.84 h (SE 0.73 h) (19). This further supports the finding that insufficient sleep causes increased levels of fatigue and sleepiness, and decreased levels of subjective alertness and neurobehavioral performance (7). It is, therefore, important from an operational perspective to determine the amount of total sleep time during break periods so as to determine the level of cognitive alertness prior to the recommencement of work. Therefore, the general aim of our research was to determine, using information obtained from duty schedules, whether the total amount of sleep obtained during layover periods between international flights is predictable. Our research specifically aimed to determine the importance of accounting for geophysical factors, such as the number of hours of a layover period occurring at local or biological night, on sleep.

**METHODS**

**Participants**

In total, 86 Qantas Airways Limited cockpit flight crew personnel consisting of captains (n = 31), first officers (n = 33), and second officers (n = 22) were included in the current analysis (mean age ± SD = 46.7 ± 4.3 yr). These crews flew either Boeing 747 or 767 aircraft on selected international patterns into and out of Australia. This group was involved in data collection as part of the Fatigue Risk Management System project currently being undertaken by Qantas, the Australian Civil Aviation Safety Authority, the Australian and International Pilots Association, and the University of South Australia’s Centre for Sleep Research. In order to eliminate results masked by random interindividual variation, no two crewmembers undertook the same flight pattern at the same time, to eliminate potential intercrew variability.

Patterns used in the analysis include: Australia-Singapore-Europe-Bangkok-Australia (n = 42); Australia-Los Angeles-Auckland-Australia (n = 19); Australia-Los Angeles-New York-Los Angeles-Australia (n = 10); and Australia-Hong Kong-Australia (n = 19), and were representative of the international flight and duty patterns operated by Qantas. One layover was randomly selected from each pattern for analysis to eliminate potential interindividual variation (Table I).

Predictions were made on a cross-validation dataset separate from the data used to formulate the models. The validation dataset chosen for this purpose was collected concurrently with the data previously described in the Methods section and separated prior to analysis. The validation dataset consisted of 11 participants flying either to Europe (n = 8), Los Angeles (n = 1), New York (n = 1), or Hong Kong (n = 1). These participants were chosen to be included in the validation set as they were part of crews which were included in the testing sample analyzed above. That is, they were excluded from the initial analysis because of possible problems of intercrew variability, which linear regression models are unable to account for. The layover lengths for this cross-validation data ranged between 25.87 and 63.77 h (mean ± SD, 52.34 ± 12.39). The average total sleep time during the layover periods was 14.04 ± 3.39 h.

**Recruitment**

Qantas aircrew personnel were informed of the proposed study by e-mail and by poster before being invited to information sessions to explain the study’s purpose and protocol in full. All individuals received an information sheet and it was explained that participation was entirely voluntary and that they were free to withdraw at any stage during the study without prejudice. Pilots were not screened for sleep disorders in the recruitment process. However, the Australian Civil Aviation Safety Authority requires all Australian commercial pilots to undergo regular examination by a
designated aviation medical examiner to assess fitness for duty. This occurs annually for pilots under the age of 40 and biannually thereafter. These examinations consist of electrocardiogram measurement, optical, auditory, blood, and urine testing. These medical exams also require the completion of a general health questionnaire including questions regarding height, weight, body mass index, and any current medications or medical conditions.

This examination also requires designated aviation medical examiners to specifically inquire whether or not applicants present with symptoms suggestive of obstructive sleep apnea. In the event that symptoms consistent with a positive diagnosis are discovered, applicants must complete an Epworth Sleepiness Scale and submit to a further examination by a sleep physician if a score greater than 16 is achieved. If sleep disturbances become evident during these further examinations, treatment must be obtained and compliance maintained for a medical certificate to be issued.

Those selected gave their informed consent and were not paid for taking part in the study other than their usual salary while at work. Based on their duty periods and patterns flown, subjects were then selected to wear activity monitors and complete sleep and work diaries. Participants were sent an activity monitor by mail and asked to wear it for a minimum of 15 d, beginning 4 d prior to departure. Sleep and work diaries were also completed during this time. The study had approval from the University of South Australia’s Human Research Ethics Committee.

Work Setting

Each recorded duty period contained at least 4 d at home (in order to minimize or eliminate existing circadian disruption) followed by at least two international flights (one outbound from and one inbound to Australia and any other international flights undertaken while in transit), with crews spending time recovering between flights during layover periods. During these layover periods flight crews were free to do as they wished and were not required to undertake any work-related duties.

During flight, crews undertook a variety of tasks usually lower in workload, except for takeoff and landing where workload is increased. While the aircraft is at cruising altitude, the members of the flight crew are given the opportunity to take in-flight naps in crew rest areas situated away from the flight deck. These naps are coordinated to ensure the flight deck is attended by either the captain or first officer at all times. All crewmembers have the opportunity to take at least one nap, possibly more, depending on the length of the flight and crew compliment.

Participants also completed a duty diary, which required them to provide information about the start and end time of each flight. This also included the Samn Perelli Fatigue scale (14), a subjective fatigue scale designed especially for air transport workers, to enable the crewmembers to rate their levels of fatigue before and after each duty period. The sleep/wake behavior of flight crew was monitored objectively using wrist activity monitors (Mini Mitter, Sunriver, OR) and Sleepwatch software (Actiware-Sleep®9©, Cambridge Neurotechnology Ltd., Cambridge, UK). The actiwatch contained a piezo-electric accelerometer that detected movement with a resultant force above 0.01 g with a sampling rate of 32 Hz. Flight crew wore the activity monitor at all times during the participation period unless showering or in situations where the monitor was likely to be damaged.

Statistical Analysis

Prior to analysis outliers were excluded (defined as more than 3 SDs from the mean). One data point was excluded using this method. The initial analysis aimed to determine which variables were highly intercorrelated with each other. The variables tested were: number of time zones crossed; layover length; prior flight duration; the number of social night hours (SNH, defined as the number of hours between 21:00 and 07:00 in local time); and the number of biological night hours (BNH, defined as the number of hours between 21:00 and 07:00 in Australian Eastern Standard Time, with this time zone used as a proxy for biological clock time). Correlations were conducted in order to create an association matrix to analyze the level of orthogonality among the dependent and independent variables.

Bivariate linear regressions were then conducted so that significant predictors of total sleep time could be found. These linear model parameter estimates were then used to predict the total sleep time during the layovers in the cross-validation dataset. Plots of actual vs. predicted total sleep time (TST) are included for the cross-validation datasets where appropriate, including linear lines of best fit. Unless stated otherwise, signifi-
Correlation analysis indicated that many variables were highly correlated with each other and with the dependent variable TST (Table II). In light of these results, five bivariate regressions were carried modeling TST against each independent variable. The r-values obtained from these regression analyses are contained in column six of Table II. The parameter estimates for each model were then used to predict the TST obtained in the cross-validation dataset described previously. The correlations of these fitted models are presented in the last column of Table II. Fig. 1 shows the five independent variables plotted against TST using the original testing dataset and the cross-validation dataset.

**RESULTS**

Correlation analysis indicated that many variables were highly correlated with each other and with the dependent variable TST (Table II). In light of these results, five bivariate regressions were carried modeling TST against each independent variable. The r-values obtained from these regression analyses are contained in column six of Table II. The parameter estimates for each model were then used to predict the TST obtained in the cross-validation dataset described previously. The correlations of these fitted models are presented in the last column of Table II. Fig. 1 shows the five independent variables plotted against TST using the original testing dataset and the cross-validation dataset.

**DISCUSSION**

The aim of this study was to examine the total amount of sleep obtained throughout layovers by international flight crews during transmeridian flight patterns. Five prospective variables were analyzed using linear correlations and bivariate regression to determine whether TST during layovers is predictable in complex transmeridian environments. Furthermore, we aimed to determine whether prospective variables could be used to obtain accurate estimates of the amount of sleep obtained during layover periods.

Correlation analysis indicated that layover length, SNH, and BNH are strongly significantly correlated. Furthermore, BNH and flight duration were also strongly correlated (p < 0.05). Interestingly, time zone displacement was highly correlated to BNH, SNH, layover length, and prior flight duration. Inspection of the data indicated that time zone displacement is highly correlated with these variables due to the scheduling practices undertaken by the airline. For example, current schedules are created such that the longer the flight duration from Australia, the longer the recovery period allowed during layovers prior to the next flight. Furthermore, the longer the flight duration from Australia, the more time zones are crossed.

Linear regression models were then constructed to determine which factors were able to significantly predict TST. Results indicate that layover length followed by SNH followed by BNH are the best predictors of TST ($r = 0.9, 0.8$, and $0.7$, respectively). These results were reinforced when the model parameters were used to estimate the TST obtained during layovers from a cross-validation dataset. The data used for cross-validation was collected from crewmembers on some of the same flights and layovers of the testing dataset. These data were initially excluded from the original model parameterization and analysis due to potential intercrew vari-

![Fig. 1](image-url) Plots showing graphs of all independent variables (biological night hours, social night hours, layover length, flight duration, and time zone displacement) vs. total sleep time for the cross-validation dataset. Linear lines of best fit are also included in each panel graph indicating the best fit for the whole cross-validation dataset.
The results of this validation showed that layover length and SNH are equally adept at predicting TST during layovers \((r = 0.83\) and \(0.85\)). Interestingly, these obtained higher results than the BNH model, which obtained a correlation coefficient of \(0.7\).

The results of this paper compare favorably with those found in similar studies in a non-transmeridian environment. Roach, Reid, and Dawson (12) found that the amount of sleep obtained during break periods was dependent on the break length and time of break onset for locomotive engineers. Not surprisingly, this study has shown that the TST obtained by flight crews is also dependent on the length of the break period. The models presented in this study have not attempted to include any explicit time of day components, as doing so is difficult in a transmeridian environment. Instead this study has used the variables social and biological night hours. It was found that the SNH model is more robust than the BNH model (test data, \(r = 0.8\) vs. \(0.7\); validation data, \(r = 0.8\) vs. \(0.7\)). This indicates that social factors play an important part in sleep during layovers. This finding supports studies already carried out in the literature; for example, Lowden and Åkerstedt found that \(90\%\) of flight crews adopt local sleep times during a layover after westward travel (8), with a similar number of participants selecting local night sleep times after eastward travel (9).

The results obtained during this analysis can only be used to give indications of factors which may be significant predictors of TST. The dataset described and used in the current analysis was not randomized in any way prior to collection to eliminate problems with confounding variables, and as such was a dataset “of convenience.” However, due to the operational objectives of the airline, it was impossible for the authors to collect data in any other way. It was impossible to obtain more evenly distributed data as the airline in question conducts certain flight patterns each week to certain destinations, and this rarely changes. Furthermore, each pattern is allocated set arrival and departure times and layover lengths in advance. One future direction of this current analysis, therefore, is to apply this methodology to a more random sample of data, perhaps from numerous airlines.

The sleep data collected and used in the current analysis were obtained using actigraphy rather than from polysomnography (PSG), the ‘gold standard’ in sleep measurement. It has been shown that actigraphy identifies sleep onset earlier than PSG (17) and has been shown to be moderately valid and reliable for differentiating between sleep and wake in normal, healthy adult populations, but becomes less reliable at this differentiation as sleep becomes more disturbed. As stated by Tryon (17), this implies that actigraphy is often unable to correctly identify sleep onset latency, underestimating the time of sleep onset and overestimating TST. However, due to the many operational limitations of this study, actigraphy was seen as a viable alternative, especially since there is recent evidence suggesting the error between actigraphy and PSG is systematic rather than random (17).

**Implications for Models of Fatigue, Alertness, and Performance**

One of the limitations of current fatigue and alertness models is their failure to accurately predict the cumulative adverse effects of chronic sleep restriction (18). After further testing and validation, one of the further applications of the TST models outlined in this paper is a future extension to predict possible excessive cumulative wakefulness or sleep deficit during layovers compared with baseline sleep at home. One interesting topic of research would, therefore, be to make predictions of the average alertness levels of flight crews based on the estimated levels of cumulative sleep debt experienced during a trip.

In conclusion, this study has provided further proof that social factors may be a better predictor of TST during layover periods than biological factors. It is clear that in the safety critical setting of aviation, it is important to be able to predict in advance the TST by cockpit flight crews. This could then allow researchers and duty schedulers to predict crew’s corresponding fatigue levels, especially at the end of a layover prior to takeoff, a critical phase of a flight.

**ACKNOWLEDGMENTS**

We thank the Qantas flight crews who volunteered to take part in this study, and Trace Sletten and David Darwent, who assisted with data collection and analysis. We would also like to thank Dr. Jill Dorrian and the three anonymous reviewers who assisted with multiple drafts of this manuscript. Their assistance has improved the quality of this paper considerably.

This work was supported by Qantas Airways Limited, the Australian Civil Aviation Safety Authority, the Australian and International Pilots Association, and the Australian Research Council.

**REFERENCES**

12. Roach GD, Reid KJ, Dawson D. The amount of sleep obtained by...