TNO Defence, Security and Safety

Kampweg 5
P.O. Box 23
3769 ZG Soesterberg
The Netherlands

www.tno.nl

T +31 346 35 62 11
F +31 346 35 39 77
Info-DenV@tno.nl

TNO report

TNO-DV 2008 C512

Sleep, fatigue, and alertness in North Sea Helicopter Operations

Date	December 2008
Author(s)	M. Simons, MD E. Wilschut, MSc P.J.L. Valk, MSc
Assignor	Inspectie Verkeer en Waterstaat (IVW), Toezichteenheid Luchtvaart Operationele Bedrijven (TE LOB)
Project number	032.10184
Classification report	Unclassified
Title	Unclassified
Abstract	Unclassified
Report text	Unclassified
Appendices	Unclassified
Number of copies	10
Number of pages	35 (incl. appendices, excl. distribution list)
Number of appendices	4

Executive summary

Title	:	Sleep, fatigue, and alertness in North Sea Helicopter Operations
Author(s)		M. Simons, MD E. Wilschut, MSc P.J.L. Valk, MSc
Date		December 2008
Reportnr.	:	TNO-DV 2008 C512

Introduction and aim of the study

Almost all North Sea Helicopter operations from Den Helder Airport (De Kooy) involve flights to offshore platforms in the North Sea. Most flights involve multiple sectors to different platforms. Flights are characterised by multiple landings, with an average of just 17 minutes between take-off and landing. Take-offs and landings are often made under difficult weather, lighting, and turbulence conditions. Under these conditions, fatigue and impaired fitness may lead to lowered levels of alertness and performance during critical phases of flight. Prevailing Flight and Duty Time Limitations (FTLs) are derived from fixed-wing operations and may insufficiently apply to the specific aspects of North Sea helicopter Operations. The potential problem area has been identified by the Transport and Watermanagement Inspectorate and the sector. Ministry of Transport, Public Works and Watermanagement, Directorate General for Aviation and Maritime Affairs (DGLM) commissioned TNO Defence, Security, and Safety (department Human Performance) to conduct 1) a literature study on workload and fitness of North Sea helicopter pilots in relation to safety and health aspects, and 2) a field study on workload, fatigue, and alertness of pilots involved in these operations.

The first phase of the study (literature study) has been reported separately in TNO report TNO-DV 2008 C027. In summary this report had the following conclusions:

- Sufficient pre-duty sleep is an important prerequisite for sufficient in-flight alertness. Based on prevailing work rosters in Dutch North Sea helicopter operations, it was supposed that impaired and cumulative sleep debt might possibly occur in these operations.
- Turbulence around offshore platforms, unfavourable weather conditions, distraction by calculating load/fuel or filling in paper work during the flight, timeliness of information received, and helideck lighting are the most important factors that can affect workload and the safety risk. Other factors that may affect fatigue are the number of landings, irregular duties, reporting for duty during the WOCL, late reporting off duty, work rosters, wearing exposure suits, light conditions, and cabin environmental conditions, such as vibrations, ventilation, and temperature. All mentioned factors may affect fatigue and alertness levels and consequently flight safety. The safety risk may further be affected cockpit misting, and the proficiency level of the Helicopter Landing Officer.
- The prevalence of acute and chronic back pain and disorders is significantly higher in helicopter pilots than in fixed-wing pilots, or non-flying control groups.
  Acute back pain may increase the safety risk, when it occurs during a flight.

Present report

The present report concerns the second phase of the study (2): the field study. The aim of the present study was to assess the effects of workload and fatigue on alertness-related flight safety of pilots engaged in Dutch North Sea helicopter operations and, if necessary, to make recommendations to improve flight safety. Assessments were to be performed during 1) a period with favourable light conditions (summer), and 2) a period in which light conditions are more unfavourable (winter).

Method

All pilots of CHC Helicopters Netherlands (Canadian Helicopter Corporation) and Bristow Helicopters, the two main operators of Dutch North Sea helicopter operations, were asked to participate on a voluntary and unpaid basis. Confidentiality and anonymity of subject's data were guaranteed. Pilots were regularly scheduled in their normal duty rosters in two 14-days assessment periods (Summer and Winter). Each duty day a maximum of 5 test sessions had to be performed: 1) after wake up, 2) before start of duty (pre-duty), 3) at a quiet moment halfway the duty, 4) at the end of duty (end-duty), and 5) at bedtime. On days off, 2 test sessions had to be performed: after wake up and at bedtime. Each test session was performed on a PDA and took approximately 8-10 minutes. Test sessions involved completion of several questionnaires related to operational conditions and to sleep, together with validated rating scales indicative for alertness and fatigue, such as the Global Vigor and Affect scale, the Stanford Sleepiness Scale, and the Need for Recovery scale. In each session, a 5-min vigilance task (VigTrack) was performed to provide objective data. To objectify data on sleep, volunteers wore an Actiwatch device during sleep at home.

Study Outcome

Description of results

Data sets of CHC and Bristow pilots were analysed separately and will be described separately.

CHC Results

Data sets of twenty-eight 14-days assessment periods, of which 15 were summer assessments (May-June, 2007) and 13 winter assessments (January-March, 2008), were included in the analysis. A total of 24 pilots (mean total flight hours: 3602; range 290-11000) participated; 21 were male and 3 were female (mean age 37 yrs; range: 27-58). Participants were 11 captains and 13 first officers. Of the four helicopter types used by CHC, 46% of pilots flew most frequently S76, 29% AW139, 17% EC155, and 8% reported to regularly fly S61. Participants were reasonably good sleepers. Mean travel time from home to airport De Kooy was 54 min (range 5-125). The mean reported daily Flight Duty Period (FDP) in the study periods was 6:28 hr (range 2:15-10:45), the earliest reporting time was 06:00 h, while the latest reporting off duty was at 22:45 h.

Pre-Duty sleep at home

The mean pre-duty total sleep time (TST) was 6:44 h subjectively and 5:38 h objectively and sleep quality was good. Pre-duty sleep had a slightly poorer quality and a shorter subjective, as well as objective, mean TST than sleep on days off. Before duty days pilots went earlier to bed and woke up earlier than on days off. Pilots with duties starting before 12:00 h (noon) had poorer pre-duty sleep quality and shorter sleep than pilots with duties starting after 12:00 h. The difference in TST was 1:37 hr subjectively and 0:57 hr objectively. As there is sufficient scientific evidence that the length of pre-duty sleep is an important determinant of pre-flight and in-flight alertness, we consider that a mean pre-duty TST of 6 hrs or less, as is found in pilots starting their duty in the morning, may lead to undesirable levels of alertness in some cases. Therefore, we emphasize the importance of sufficient pre-duty sleep.

Travel time to the airport (De Kooy)

Longer travel time from home to De Kooy was associated with shorter pre-duty sleep and was weakly correlated with earlier bedtime, poorer sleep quality, and more fragmentation of sleep. Longer pre-duty travel time was also correlated with lower levels of vigilance performance and higher sleepiness levels before and during the duty, although values never approached a risk zone during the FDP.

Workload, fatigue, alertness, and vigilance during FDPs
Higher experienced workload was associated with longer duty duration, more flight hours, and a higher number of landings. Worse weather conditions were correlated with more effort to perform the landings. More effort experienced to perform the landings was associated with higher experienced workload. Flying the AW139 helicopter was associated with lower scores on work demand, workload, and effort to perform the landings, compared with the S76 and EC155. The S61was not included in the statistical analysis of comparisons between helicopter types, because it was flown by only very few participants.

Highest sleepiness levels and lowest vigor scores were found at bedtime, which is a normal phenomenon, caused by fatiguing effects of a working day in combination with sleep pressure dictated by the circadian clock.

During the FDP, alertness and (objective) vigilance levels were always high and never approached risk levels associated with alertness-related flight safety. Mean pre-duty, halfway-duty, and end-duty sleepiness scores (inverse of alertness) were never higher than 2.3, where level 2 stands for 'functioning at a high level, but not at peak; able to concentrate' and level 3 signifies 'relaxed; awake; not at full alertness; responsive'. As level 4 ('a little foggy; not at peak; let down') can be considered as the first level where alertness-related flight safety approaches the risk zone, it can be concluded that alertness-related safety was always maintained at a safe level during the FDPs. Fatigue levels, indicated by vigor, sleepiness, and 'need for recovery' scores never exceeded levels that are known to negatively influence safety or health.

Summer versus Winter Operations
Summer operations showed no significant differences compared with winter operations in terms of experienced workload, how demanding work was, or how much effort it had cost. Only minor differences between summer and winter weather conditions were reported and this may explain why the (potential) differences were small to absent. Vigor scores at bedtime were significantly lower in winter. However, lower vigor scores, indicating lower levels of energy and alertness, at bedtime in winter cannot be solely attributed to a higher workload on winter duty days, because bedtime vigor scores were also significantly lower on days off. Both on days off and on duty days, pilots slept longer and better in winter than in summer, which may be explained by a normal human seasonality phenomenon.

Bristow results
Due to logistic problems only five Bristow pilots participated in one study period (October 2007). As a consequence, the amount of data was too small to justify detailed statistical analysis. Based on a descriptive analysis, it was concluded that vigor, sleepiness, and tracking performance scores of the five participating Bristow pilots indicate that alertness and fatigue levels during the FDP never approached risk levels associated with alertness-related flight safety. The mean number of landings per FDP was considerable smaller than found in CHC operations, which may indicate that –at least for the participating pilots- Bristow operations differed from CHC operations. It was also concluded that mean weather conditions in the Bristow study period were less favourable compared to the study periods of CHC.

Conclusions and Recommendations

Conclusions and recommendations concerning CHC North Sea Operations

1 We conclude that during the present North Sea helicopter FDPs, pilots always maintained alertness and vigilance levels that are considered safe in terms of alertness-related flight safety. We also conclude that fatigue levels never exceeded levels that are known to negatively influence safety or health.

2 Pilots who had to report for an early duty had considerably shorter pre-duty sleep than others. Shorter sleep was associated with higher sleepiness levels and lower vigilance scores during the FDPs, although levels of sleepiness and vigilance never approached the risk zone associated with alertness-related safety. It is recommended to pursue an increase of the total pre-duty sleep time in pilots scheduled on morning duties. This may be done by stimulating awareness, both of pilots and management, of the importance of sufficient sleep and to try and guarantee sufficient pre-duty sleep opportunities, particularly before an early duty.

3 Pilots with longer travel time to airport De Kooy had lower levels of vigilance and higher sleepiness levels during and after the FDP. These pilots also had shorter pre-duty sleep, which may signify the causal relation with lower vigilance and higher sleepiness. Although in these cases, vigilance and sleepiness values never approached a risk zone of alertness-related safety, it is recommended to stimulate opportunities for shorter travel time from home to De Kooy, whenever socially acceptable and possible.

4 Differences between summer and winter operations were minimal in terms of workload, vigilance, and sleepiness scores. Mean weather conditions were experienced as only slightly worse during the summer, which may explain the lack of differences between both seasons. Both on days off and on duty days, pilots slept longer and better in winter than in summer and had lower bedtime vigor scores in winter. Both phenomena are considered as normal human seasonality effects.

5 Flying the AW139 helicopter was associated with lower scores on work demand, workload, and effort to perform the landings, compared with the S76 and EC155.

Conclusions concerning Bristow North Sea Operations

Vigor, sleepiness, and tracking performance scores of the five participating Bristow pilots indicate that alertness and fatigue levels during the FDP never approached risk levels concerning alertness-related flight safety. The small amount of data did not allow for more detailed conclusions, such as those described in the CHC part of this section.

Overall conclusion of the project (phase 1 and 2)

Taking into account the conclusions of the report of phase 1 (TNO DV 2008 C027) and above-mentioned conclusions, we conclude that flight safety in North Sea helicopter operations may be primarily affected by other factors than low alertness or high fatigue levels. In this context, principal factors are turbulence around offshore platforms, unfavourable weather conditions, distraction by calculating load/fuel or filling in paper work, and poor helideck lighting. Additional unfavourable factors may be wearing exposure suits, light conditions, cockpit misting, and the proficiency level of the Helicopter Landing Officer. Cabin environmental conditions, such as vibrations, ventilation and temperature, and unfavourable body positions may contribute to health problems. In that context, it would be useful to assess whether the introduction of new types of helicopters, such as the EC 155 and AW 139, will reduce health problems that were always considered to be specific for helicopter pilots (e.g. back pain).

Acknowledgements

We would like to thank Bert Vos, inspector of the Transport and Watermanagement Inspectorate, and his successor inspector Wandert Brandsen for initiating us in the world of Dutch North Sea Helicopter operations and for their role in planning the study.

The yield of field studies in aviation depends highly on smooth and well-organized logistic support. Thanks to the efforts of the Canadian Helicopter Corporation (CHC) and Bristow Helicopters staffs, the present study was successfully performed.
We especially thank Dé Jansen, Flight Crew Manager CHC Helicopters Netherlands, Erik van Gelder, pilot CHC, and Adrian Rose, Chief Pilot Bristow, who organized and coordinated distribution and collection of study equipment and scheduling of the participating pilots and liaison with the researchers. We would like to thank the participating pilots for their dedication and effort to provide useful data.

Contents

1 Introduction

Almost all North Sea Helicopter operations from Den Helder Airport (De Kooy) involve flights to offshore platforms in the North Sea. In the Dutch offshore sector it is common to have a large number of platforms within a small area and most flights involve multiple sectors to different platforms. Flights are characterised by multiple landings, with an average of just 17 minutes between take-off and landing. Take-offs and landings are often made under difficult weather, lighting, and turbulence conditions. Under these conditions, fatigue and impaired fitness may lead to lowered levels of alertness and performance during critical phases of flight. Prevailing Flight and Duty Time Limitations (FTLs) are derived from fixed-wing operations and may insufficiently apply to the specific aspects of North Sea helicopter Operations. These potential problem areas have been identified by the Transport and Watermanagement Inspectorate (IVW) and the sector. In a previous consultation between TNO and the Inspectorate the following factors, that may potentially impair fatigue, alertness, and fitness, have been considered:

- Operational characteristics: shift work with consequent sleep debt, operating in the Window of Circadian Low (WOCL, between 02:00 and 06:00), high number of take-off/landings, long Flight Duty Periods (FDPs), last minute changes.
- External work environment: unfavourable weather conditions, sometimes poor landing facilities on the platforms, limited air traffic control facilities, no or limited radar coverage.
- Internal work environment: ventilation, vibrations, noise, temperature, ditching exposure suit, seats.

The ambition of the Directorate General for Aviation and Maritime Affairs (DGLM) is to improve the safety level of helicopter operations. In that context DGLM commissioned TNO Defence, Security, and Safety (department Human Performance) to conduct:

1. a literature study on workload and fitness of North Sea helicopter pilots in relation to safety and health aspects;
2. a field study on workload, fatigue, and alertness of pilots involved in these operations.

The first phase of the study (literature study) has been reported separately in TNO report TNO-DV 2008 C027 [Simons, 2008]. In summary this report had the following conclusions:
Based on the analysis of available national and international data, it was concluded that sufficient pre-duty sleep is an important condition of sufficient in-flight alertness. Based on prevailing work rosters in Dutch North Sea helicopter operations, it was supposed that impaired and cumulative sleep debt might possibly occur in these operations. It was also concluded that turbulence around offshore platforms, unfavourable weather conditions, distraction by calculating load/fuel or filling in paper work during the flight, timeliness of information received, and helideck lighting are the most important factors that can affect workload and the safety risk. Other factors that may affect fatigue are the number of landings, irregular duties, reporting for duty during the WOCL, late reporting off duty, work rosters, wearing exposure suits, light conditions, and cabin environmental conditions, such as vibrations, noise, ventilation, and temperature. All mentioned factors may affect fatigue and alertness levels and consequently flight safety. Specific relations between fatigue and radio-telephony (R/T) noise could not be determined due to a lack of data in literature. The safety risk may further be affected by factors not related to fatigue, such as cockpit misting, and the proficiency level of the Helicopter Landing Officer. It was further concluded that the prevalence of acute and chronic back pain and

disorders is significantly higher in helicopter pilots than in fixed-wing pilots, or non-flying control groups. Acute back pain may increase the safety risk, when it occurs during a flight.

The present report concerns the second phase of the study (2): the field study.

1.1 Aim of the present project – phase 2: field study

The aim of the project was to assess the effects of workload and fatigue on alertness-related flight safety of pilots engaged in Dutch North Sea helicopter operations and, if necessary, to make recommendations to improve flight safety. In this context, the present study involved assessments of the effects of pre-duty sleep, travel time to airport De Kooy, operational characteristics, workload, and season on fatigue, alertness, and vigilance during the daily flight duty periods.

It was decided that assessments were to be performed of pilots engaged in operations of the two main operators of Dutch North Sea helicopter operations, CHC Helicopters Netherlands (Canadian Helicopter Corporation) and Bristow Helicopters. Because it was considered that environmental light conditions may influence workload, it was decided to plan the assessments during 1) a period with favourable light conditions (summer), and 2) a period in which light conditions are usually more unfavourable (winter). The ambition was to study the same pilots during these two study periods.

2 Method

2.1 Subjects

CHC and Bristow Helicopters provided full support to the logistics of the study and all contacts with the pilots were maintained via the flight departments. All pilots of CHC and Bristow were asked to participate on a voluntary and unpaid basis. Confidentiality and anonymity of subject's data were guaranteed. Pilots received an information leaflet of TNO (see Appendix A). Pilots who volunteered received a 1 hour training session, led by a researcher, to familiarize with tests and questionnaires. Participating pilots were regularly scheduled in their normal duty rosters in two 14-days assessment periods (Summer and Winter).
Participants of this study included 24 pilots of CHC and 5 pilots of Bristow. Because the number of pilots working in the Dutch branch of Bristow Helicopters is much smaller than the number of pilots working for CHC, it has been difficult to include a sufficient number of Bristow pilots. This was further complicated by a mismatch between availability of pilots and availability of study equipment, because of detachment abroad and EC155 conversion training periods followed by all potentially participating Bristow pilots.

2.2 Assessment Methods

2.2.1 Actiwatch

Figure 1 Actiwatch.

During the duty days and days off, participants wore an Actiwatch device (Cambridge Neurotechnology; see picture) during their sleep at home. Using the event button, subjects marked beginning and end of their sleep periods. This method provides objective data on pre-duty sleep at home. The Actiwatch measures activity by means of a piezo-electric accelerometer, which records the integration of intensity, amount and duration of movements in all directions. The activity data are converted and stored in the memory unit of the Actiwatch and, after collection, downloaded to a PC through a reader interface and processed and analyzed with dedicated software (Actiwatch Sleep & Activity Software V 5.32, Cambridge Neurotechnology). Actigraphy is a generally accepted method to objectively assess sleep characteristics, such as total sleep time, activity, and fragmentation [e.g. Sadeh et al., 1989; Ancoli-Israel et al., 2003].

2.2.2 *Vigilance Task (VigTrack)*

Figure 2 Vigilance and Tracking test.

The Vigilance and Tracking test (VigTrack) is a dual-task measuring vigilance performance under the continuous load of a compensatory tracking task. The task has originally been developed on a Psion 3a palmtop computer [Valk et al., 1997]. In the present study, the VigTrack version specifically designed for use on a PDA was used (see picture; instruction Annex 1).The test is self-administered and needs circa 20 minutes practice time to eliminate significant learning effects [Valk et al., 1997]. The Vigilance and Tracking task was identified as a sensitive task tapping vigilance performance during flight operations. The task has been successfully applied in field studies concerning effects of fatigue and sleepiness in pilots [Valk & Simons, 1998; Simons & Valk, 1998, Valk et al., 2003] and was successfully applied in laboratory studies to demonstrate detrimental (residual) effects of alcohol, sedative effects of antihistamines as well as residual effects of hypnotics under conditions of simulated cabin pressure in a hypobaric chamber [Valk et al., 1997; Valk et al., 2004; Simons et al., 2006]. In the present study, the aim was to assess vigilance performance of the pilots during their Flight Duty Period (FDP).

2.2.3 *Questionnaires/rating scales*

At the start of the study pilots completed a personal demographic questionnaire (age, gender, total flight hours logged, crew assignment, etcetera)

In each test session, pilots rated their levels of vigor and alertness using the Global Vigor and Global Affect visual analogue scale [GVA; Monk, 1989] and levels of sleepiness using the Stanford Sleepiness Scale [SSS; Hoddes et al., 1973]. Details of the GVA are shown in Appendix B. The GVA is a well-validated scale, of which the vigor part (GV) is particularly relevant to assess vigor and alertness aspects [Monk, 1989] of pilots during their FDP. The scale ranges from 0 to 100, where lower values represent lower levels of alertness and vigor (‘vitality’).

Details of the interpretation of the SSS are shown in Appendix C. The SSS was used to assess subjective sleepiness throughout the FDP. The scale ranges from 1 to 7, where lower values represent lower sleepiness levels (equivalent to higher alertness levels). This subjective rating scale has proven to be sensitive in detecting any significant increase in sleepiness or fatigue [Simons et al., 1994; Simons & Valk, 1997; Simons & Valk, 1998; Valk et al., 2001; Valk et al., 2004]. SSS ratings showed to be highly correlated with flying performance and threshold of information processing speed during periods of intense fatigue [Samn & Perelli, 1982].

Fatigue after a duty day was indirectly assessed using the Need for Recovery scale (NFR; details shown in Appendix D). The NFR is a short, simple, but adequate measure for early symptoms of fatigue at work, for use in both health surveillance and scientific research [van Veldhoven & Broersen, 2003; Jansen et al., 2003]. Correlations between the NFR and other scales measuring fatigue at work are all above 0.65 [van Veldhoven & Broersen, 2003].

The quality of pre-duty sleep (at home) and during the days off was assessed by means of the Groningen Sleep Quality Scale [GSQS; Meijman et al., 1987]. Scores range from 0-14, where higher scores indicate poorer sleep quality. The GSQS has been used in a variety of studies on sleep disturbances among Dutch airline pilots [Simons et al., 1994; Valk & Simons, 1998; Simons & Valk, 1998; Valk et al., 2003]. Furthermore, subjects had to record the subjectively estimated bedtime, wake-up time, getup time, total sleep time, and number of awakenings.

2.2.4 *Test sessions*

Each work day a maximum of 5 test sessions had to be performed:

1 After wake up.
2 Before start of duty (pre-duty).
3 At a quiet moment halfway the duty.
4 At the end of duty (end-duty).
5 At bedtime.

A Pre-Duty session was only to be completed when the time period between wake up and reporting for duty was more than 2 hrs. Because it was considered not feasible, no test sessions were planned during the flight. Instead, to assess alertness levels and vigilance performance during the FDP, a test session was planned halfway duty. Pilots were asked to plan this session approximately halfway their duty, when they had a quiet period of at least 10 minutes to complete the session (e.g. during waiting at De Kooy). Each test session took approximately 10 minutes. On days off, 2 test sessions had to be performed: after wake up and at bedtime.

Each test session included at least completion of the GVA and SSS ratings and performance on the VigTrack task (5 min). This basic test session took approximately 7 minutes. At certain test sessions extra ratings, relevant for the specific session, were completed (taking an extra of 3 minutes). Each wake-up session, questions concerning pre-duty sleep were completed together with the GSQS sleep quality rating scale.
At post duty sessions, pilots completed questions concerning their duty (start, end), characteristics of workload, number of landings, and questions concerning weather conditions. Test sessions are schematically presented in Table 1.

Table 1 Schedule of test sessions.

Session	Day Off	Duty Day
Wake up (30-60 min after)	GVA, SSS, VigTrack, GSQS, sleep variables	GVA, SSS, VigTrack, GSQS, sleep variables
Pre-Duty* (at reporting)		GVA, SSS, VigTrack
Halfway Duty		GVA, SSS, VigTrack
End-Duty (at reporting off)		GVA, SSS, VigTrack number landings, duty times, weather, workload
Bedtime	GVA, SSS, VigTrack	GVA, SSS, VigTrack, NFR

* Pre-Duty session only to be completed when >2 hrs between wake up and reporting for duty.

2.3 Procedure

All test sessions were self-administered and performed on a Personal Digital Assistant (PDA). Pilots received the PDA together with an Actiwatch (see Assessment Methods) and a brief paper instruction (see Appendix A) before the start of the 14-day study period. The equipment was collected after the end their study period by the Crew Coordination Unit. Test sessions had to be performed on days off, and on regular duty days. They also had to wear an Actiwatch to record sleep at home. After collection of the PDA and Actiwatch, the results were downloaded by a researcher.

2.4 Statistical analyses

Using Statistica Data Analysis Software (StatSoft®), all variables were analysed using descriptive techniques. Those performance variables that were considered to be interesting for further analysis were tested using Student's-*t*. Subjective ratings were analysed using non-parametric techniques (Wilcoxon Matched-Pairs Signed-Ranks, Mann-Whitney U). Relationships between different variables and methods were investigated by using correlational computations (Pearson product-moment or Spearman Rank correlation coefficients).

2.4.1 Description of results

Data sets of CHC and Bristow pilots were analysed separately. Results concerning CHC operations are described in Paragraph 3 (Sections 3.1-3.5). Results of Bristow pilots in Paragraph 4 (Sections 4.1-4.4).

3 CHC: Results

Sections 3.1-3.4 describe the results of all assessments performed by CHC pilots, summer and winter combined, while Section 3.5 describes the differences between the results obtained in summer and winter.

3.1 Subjects

Data sets of twenty-eight 14-days assessment periods, of which 15 were summer assessments (May-June, 2007) and 13 winter assessments (January-March, 2008), were included in the analysis. A total of 24 pilots participated; 21 were male and 3 were female. Participants were 11 captains and 13 first officers. Of the four helicopter types used by CHC, 46% of pilots flew most frequently S76, 29% AW139, 17% EC155, and 8% reported to regularly fly S61.

Eight pilots considered themselves as an ‘evening’ type, 5 as ‘morning’ type, while 11 had no preference for evening or morning (indifferent). Further characteristics of the participants are presented in Table 2.

Table 2 Characteristics of participating CHC pilots. The last four questions relate to experience of the pilots in the last 6 months preceding the study. Average sleep quality on days off and on duty days represents scores on a Visual Analogue Scale (VAS) ranging from 0 (very poor) to 100 (excellent). N=24.

Question	Mean	Range
Age	37 yrs	27 - 58
Total flight hours logged	3602 hrs	290 - 11000
Travel time from home to De Kooy	54 min	5 - 125
Estimated total sleep time on days off	7:54 hr:min	5:00 - 10:00
Average sleep quality on days off	72	27 - 98
Estimated total sleep time on duty days	6:33 hr:min	4:00 – 9:00
Average sleep quality on duty days	65	35 - 98

3.2 Flight Duty Periods (FDP)

Pilots had to report the start and the end of their FDP; calculation of the time interval between both reported points in time provided the length of the FDP. The mean reported FDP in the study periods was 6:28 hr (range 2:15-10:45), on average the FDP started at 10:00 h (range 06:00-18:15) and ended at 16:30 h (range 08:45-22:45).

3.3 Sleep at home

Table 3 shows the sleep characteristics of pre-duty sleep at home and sleep on days off. Sleep before duty days had a slightly, but significantly, poorer quality (GSQS; $p < 0.05$) and a shorter subjective, as well as objective, mean total sleep time ($p < 0.00001$) than sleep on days off. Before duty days pilots went earlier to bed ($p < 0.00001$) and woke up earlier ($p < 0.00001$) than on days off.

Table 3 Sleep quality (GSQS), bedtime, wake up time, and subjective total sleep time (TST) and objective TST on days off, and sleep on duty days. Means and range ( ). Lower GSQS scores signify better sleep quality. Times in hr:min.

	GSQS	Bedtime	Wake up time	TST subjective	TST actigraphy
Day off	2.4 (0-13)	00:13 (22:00-03:30)	08:12 (05:30-11:00)	7:48 (5:10-10:30)	6:32 (3:33-9:53)
Duty Day	3.1 (0-13)	23:33 (19:45-03:00)	06:45 (03:45-10:10)	6:44 (3:55-10:45)	5:38 (0:55-8:46)

On the first day off after a duty day, pilots woke up earlier ($p < 0.05$) and had shorter subjective ($p < 0.05$) and objective ($p < 0.01$) total sleep times than on other days off. Sleep characteristics on the last day off before the start of consecutive duty days showed no significant differences with sleep on other days off.
Higher age was very weakly, but statistically significant, associated with shorter objective total sleep time ($r = 0.15$; $p < 0.05$).

Frequency distribution of reporting times showed 2 categories: duties starting in the morning (125 cases) and duties starting in the afternoon (99 cases). Pilots with duties starting in the morning (before noon) had poorer pre-duty sleep quality ($p < 0.01$) and shorter sleep ($p < 0.000001$) than pilots with duties starting after noon. The difference in total sleep time was 1:37 hr subjectively and 0:57 hr objectively (actigraphy).

Longer travel time from home to De Kooy was weakly, but significantly, correlated with earlier bedtime ($r = 0.31$; $p < 0.05$), poorer sleep quality ($r = 0.20$; $p < 0.05$) and more fragmentation of sleep ($r = 0.29$; $p < 0.05$).

3.4 Operational conditions, Alertness, and Vigilance performance

Ratings of operational conditions during duties are presented in Table 4. Vigor (GV) scores are also indicative for the level of alertness (i.e. higher scores indicate higher levels of alertness) and sleepiness scores (SSS) represent the inverse of alertness scores (i.e. higher SSS scores represent lower levels of alertness). GV and SSS scores rated during duty days are presented in Figure 3. Objective scores of vigilance (VigTrack task) are presented in Figure 4. All subjective and objective scores, including the scores for the days off, are shown in Table 5.

Table 4 Ratings (Visual Analogue Scale 0 to 100) of operational conditions during duties. Means of summer + winter combined and ±standard deviation (SD) and means of summer and winter. N=224.

Question	Mean (±SD)	Summer	Winter
Number of landings performed during the FDP	10 (6.4)	9.3	10.8
Overall weather conditions during the FDP VAS: 0 (very bad) to 100 (very good)	53 (28)	50	57
How demanding was the current duty day VAS: 0 (not demanding) to 100 (very demanding)	44 (23)	44	45
Effort to perform all landings VAS: 0 (extreme effort) to 100 (absolutely no effort)	60 (25)	59	61
Rating of overall workload during the FDP VAS: 0 (very low) to 100 (very high)	43 (21)	43	43

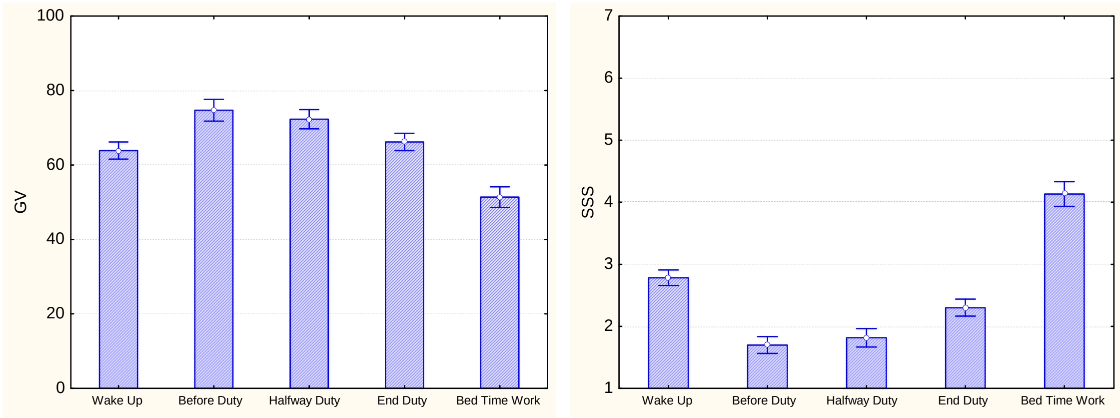

Figure 3 Ratings of vigor/alertness (GV-left panel) and sleepiness (SSS-right panel) on duty days. Means and ± 0.95 confidence interval. Higher GV scores signify higher alertness levels, higher SSS scores higher sleepiness (lower alertness).

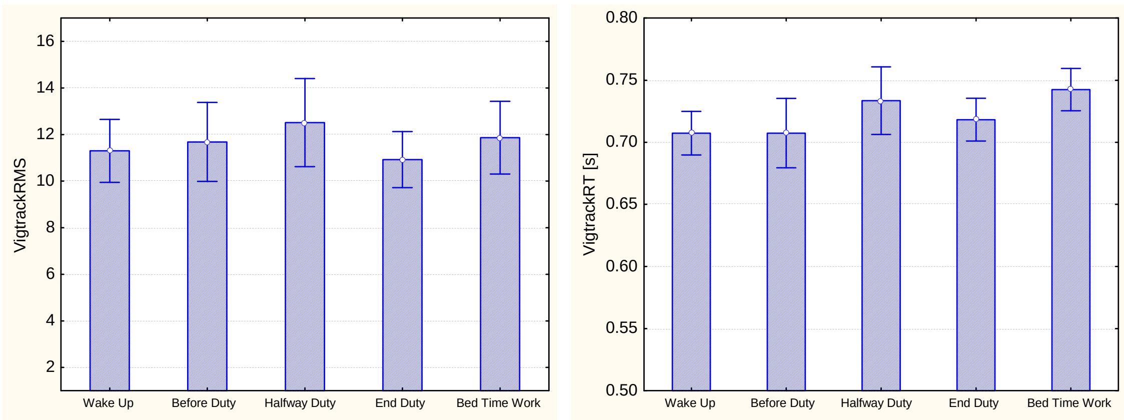

Figure 4 Scores on the VigTrack task: VigRMS (Root Mean Square of tracking error – left panel), VigRT (Reaction Time in sec – right panel) on duty days. Means and ± 0.95 confidence interval. Higher VigRMS, or RT signify lower levels of vigilance.

Table 5 Ratings of vigor/alertness (GV), sleepiness (SSS), and scores on the VigTrack task: VigRMS (Root Mean Square of tracking error), VigRT (Reaction Time in sec) on days off and duty days. Means and ±SD ( ). Higher GV scores signify higher alertness levels, higher SSS scores higher sleepiness (lower alertness). Higher VigRMS, or RT signify lower levels of vigilance.

	GV	SSS	VigRMS	VigRT
Day off				
Wake up	68.0 (18.8)	2.6 (0.9)	11.4 (6.5)	0.707 (0.081)
Bedtime	58.8 (20.4)	3.8 (1.5)	12.7 (11.1)	0.744 (0.086)
Duty Day				
Wake up	63.9 (17.5)	2.8 (0.9)	11.3 (8.1)	0.707 (0.087)
Pre-Duty	74.7 (16.2)	1.7 (0.7)	11.7 (7.7)	0.707 (0.095)
Halfway-Duty	72.3 (15.2)	1.8 (0.9)	12.5 (9.2)	0.734 (0.110)
End-Duty	66.2 (17.6)	2.3 (1.0)	10.9 (7.1)	0.718 (0.082)
Bedtime	51.4 (20.6)	4.1 (1.5)	11.7 (9.5)	0.742 (0.083)

On duty days and days off, bedtime vigor was significantly lower ($p < 0.001$) and bedtime sleepiness levels were higher ($p < 0.001$) compared to all other sessions. At pre-duty sessions vigor was highest ($p < 0.001$) and sleepiness scores were lowest ($p < 0.001$)

compared to other sessions. No significant difference was found in vigor and sleepiness levels between pre-duty and halfway-duty sessions. Vigor scores tended to be reduced at end-duty sessions to the level of wake-up sessions. The end-duty sleepiness scores were slightly higher than halfway-duty scores, but still lower than the levels after wake-up. No significant differences between pre-duty, halfway-duty, and end-duty sessions were found in vigilance performance.

Higher age was significantly ($p < 0.05$) correlated with slightly higher vigor scores ($r = 0.26$) and larger tracking error and longer reaction time on the vigilance task (VigRMS: $r = 0.37$ and VigRT: $r = 0.61$). There were no statistical differences in vigor, sleepiness, and vigilance performance between morning types and evening types. Both evening types and morning types scored higher on vigor than those who had no preference for morning or evening ('indifferent' type) ($p < 0.01$ and $p < 0.05$ respectively).

3.4.1 Pre-duty sleep

Poorer quality of pre-duty sleep (GSQS) was significantly ($p < 0.05$) correlated with lower vigor scores (GV: $r = 0.41$), higher sleepiness levels (SSS: $r = 0.45$), and worse vigilance performance (VigRMS: $r = 0.21$; VigRT: $r = 0.33$). Shorter pre-duty sleep length (TST) was significantly ($p < 0.05$) correlated with lower vigor scores ($r = 0.24$), higher sleepiness levels ($r = -0.27$), and longer reaction times on the vigilance task ($r = -0.27$).

3.4.2 Travel time from home to De Kooy

Longer travel time from home to De Kooy was significantly ($p < 0.05$) correlated with larger tracking errors (Vig RMS: $r = 0.52$) and longer reaction times (VigRT: $r = 0.70$) in the vigilance tasks performed before duty, halfway duty ($r = 0.53$ and $r = 0.55$), and at the end of duty ($r = 0.50$ and $r = 0.37$), and at bedtime ($r = 0.54$ and $r = 0.50$). Correlations between travel time to De Kooy and sleepiness scores (SSS) halfway-duty, end-duty, and at bedtime were also statistically significant, but correlations coefficients were on a much lower level (range of r: .14 to .26).

3.4.3 Morning and afternoon duties

Measurements at the end of the duty showed that pilots with duties starting in the morning (before noon) had significant higher sleepiness levels ($p < 0.05$) and larger tracking error on the vigilance task ($p < 0.0001$) than pilots with duties starting in the afternoon. There were no significant correlations between the time of day when duties ended and vigor, sleepiness, or vigilance levels.

3.4.4 Workload factors

A higher number of flight hours was weakly, but significantly ($p < 0.05$), correlated with worse vigilance performance halfway-duty ($r = 0.31$), at the end of duty ($r = 0.22$), and at bedtime ($r = 0.25$). A higher number of flight hours was also significantly correlated with higher sleepiness levels ($r = 0.26$) and with lower vigor scores ($r = -.23$) at the end of the duty, but correlation coefficients were low.

The more demanding the work was, the higher sleepiness levels and the lower vigor scores were at halfway-duty (respectively $r = 0.19$ and $r = -0.31$; $p < 0.05$) and at the end of duty (respect. $r = 0.30$ and $r = -.27$; $p < 0.05$). Higher estimated workload was weakly, but significantly ($p < 0.05$) associated with higher sleepiness levels at the end of duty ($r = 0.22$) and with lower vigor scores halfway duty ($r = -0.26$) and at the end of the duty ($r = -0.27$). The more efforts landings had cost, the slightly larger the tracking error on the vigilance task was at halfway-duty ($r = 0.26$; $p < 0.05$).

There were no significant correlations between the number of landings, or the weather conditions and sleepiness levels, or vigilance performance.

The experienced work demand was highly correlated with the experienced workload ($r = 0.88$; $p < 0.05$) and duty duration, number of landings, and number of flight hours were highly interrelated. Higher experienced workload was associated with longer duty duration ($r = 0.26$), more flight hours ($r = 0.27$), and a higher number of landings ($r = 0.32$). The later a duty ended, the higher experienced work demand was ($r = 0.22$). Worse weather conditions were significantly correlated with more effort to perform the landings ($r = 0.36$). The more effort was experienced to perform the landings, the higher the workload was rated ($r = -0.42$).

As the S61 helicopter was flown by only very few participants, this helicopter type was not included in a statistical analysis concerning differences between the various types of helicopter. There were no significant differences between the S76, AW139, and EC155 helicopters in terms of vigor, sleepiness, and vigilance levels at pre-duty, halfway duty, end-duty, and bedtime sessions. No significant differences could be shown between these types of helicopters regarding the number of landings, flight hours, or weather conditions. However, flying the AW139 was experienced as significantly less demanding than flying the S76 ($p < 0.05$), or the EC155 ($p < 0.01$), required less effort to perform the landings than the S76 ($p < 0.0001$), or EC155 ($p < 0.001$) and had lower workload ratings than S76 ($p < 0.01$), or EC155 ($p < 0.001$). There were no significant differences on these items between S76 and EC155.

The Need for Recovery (NFR) at bedtime after a duty day was 27 on a scale of 0 (no need for recovery) to 100 (extreme need for recovery). Higher need for recovery was weakly, but significantly ($p < 0.05$), correlated with earlier reporting times ($r = -0.22$), more demanding work ($r = 0.35$), higher overall workload ($r = 0.24$), and higher sleepiness levels ($r = 0.34$).

3.4.5 *Correlations between vigor, sleepiness, and vigilance scores*

Correlations coefficients between overall vigor ratings, sleepiness ratings, and vigilance performance scores are presented in table 6. Pilots with higher vigor scores had significantly lower sleepiness levels and shorter reaction times. Higher sleepiness ratings were associated with larger tracking error and longer reaction times on the vigilance task, while larger tracking error was correlated with longer reaction times.

Table 6 Correlation coefficients between vigor ratings, sleepiness ratings, and vigilance performance scores. VigRMS=root mean square of tracking error; VigRT=reaction time. Presented correlations are statistically significant ($p < 0.05$). Correlation coefficients under 0.40 are considered as low, while values between 0.41 and 0.61 are considered as moderate correlations.

	GV	SSS	VigRMS	VigRT
Global Vigor (GV)		-0.55	ns	-0.20
Sleepiness (SSS)	-0.55		0.22	0.41
VigRMS	ns	0.22		0.61
VigRT	-0.20	0.41	0.61	

3.5 Differences between Summer and Winter

There were no statistically significant differences between summer and winter in terms of experienced workload, how demanding work was, or how much effort it had cost. Table 4 shows that weather conditions were experienced as slightly worse during the summer (trend: $p < 0.075$) and the mean number of landings was slightly higher during winter (trend: $p < 0.070$)

On days off, mean sleep length was 31 minutes longer in winter than in summer ($p < 0.05$), while sleep efficiency (total sleep time / time in bed x 100%) was higher ($p < 0.05$). On duty days, mean pre-duty sleep length was 1 hour longer in winter ($p < 0.000001$), and sleep efficiency was also higher ($p < 0.05$), while sleep was slightly less fragmented in winter ($p < 0.01$).

Comparing summer with winter duty days, there were no significant differences of vigor, sleepiness, and vigilance scores in the wake-up, halfway duty, and end-duty assessments. Bedtime assessments both on duty days and days off showed higher vigor scores in summer than in winter ($p < 0.000001$).

The Need for Recovery (NFR) scores showed no significant differences between summer and winter.

4 Bristow: Results

4.1 Subjects

Data sets of five 14-days assessment periods were included in the analysis. All 5 data sets were collected in October 2007. Two incomplete data sets were collected in the second study period (April 2008). Due to missing data these two data sets were not included in the analysis. A total of 5 pilots participated, while 4 of them completed the general questionnaire. These four pilots were all male and involved 3 captains and 1 first officer. Two pilots considered themselves to be a ‘morning’ type and one considered himself to be an ‘evening’ type, one pilot indicated to have no preference. All participating Bristow pilots reported to fly the EC155 helicopter. Further characteristics of the participants are presented in Table 7. The one pilot, who did not complete the general questionnaire, did complete the operational study assessments. Therefore, results concerning the study periods are based on analysis of 5 pilots.

Table 7 Characteristics of participating Bristow pilots. The last four questions relate to experience of the pilots in the last 6 months preceding the study. N=4.

Question	Mean	Range
Age	39 yrs	27 - 49
Total flight hours logged	4500 hrs	2600 - 7800
Total flight hours logged on EC155	150	100-200
Travel time from home to De Kooy	54 min	30 - 120
Estimated total sleep time on days off	8 hrs	7:00 - 10:00
Average sleep quality on days off (% of excellent)	73%	56 - 84
Estimated total sleep time on duty days	7 hrs	7:00 – 7:00
Average sleep quality on duty days (% of excellent)	66%	42 - 83

4.2 Flight Duty Periods (FDP)

The mean reported FDP in the study periods was 7:40 hr (range 6:00-10:00), on average the FDP started at 9:23 hr (6:00-13:45 hr) and ended at 18:46 hr (range 12:00- 21:30 hr).

4.3 Sleep at home

Table 8 shows the sleep characteristics of pre-duty sleep at home and sleep on days off. Data are descriptive. Due to the small number of participants further statistical testing was considered not justified. Sleep before duty days had a slightly, poorer quality and a shorter subjective, as well as objective, mean total sleep time than sleep on days off. Before duty days pilots went earlier to bed and woke up earlier than on days off.

Table 8 Sleep quality (GSQS), bedtime, wake up time, and subjective total sleep time (TST) and objective TST on days off, and sleep on duty days. Means and range ( ). Lower GSQS scores signify better sleep quality. Times in hr:min.

	GSQS	Bedtime	Wake up time	TST subjective	TST actigraphy
Day off	3.3 (0-13)	23:52 (22:30-01:30)	07:24 (05:01-10:00)	7:34 (2:47-10:29)	6:29 (2:27-9:17)
Duty Day	4.0 (0-13)	23:19 (21:15-01:00)	6:25 (04:09-09:30)	6:19 (4:40-8:07)	5:24 (4:09-6:27)

Due to the small data sample, no relevant correlations of sleep characteristics with reporting time, duty schedules, or travel time from home to airport De Kooy could be determined.

4.4 Operational conditions, Alertness, and Vigilance performance

Ratings of operational conditions during duties are presented in Table 9.
All vigor, sleepiness, and vigilance scores, including the scores for the days off, are shown in Table 10.

Table 9 Ratings (Visual Analogue Scale 0 to 100) of operational conditions during duties. Means applying to the study period in October 2007 and ±standard deviation (SD).

Question	Mean (±SD)
Number of landings performed during the FDP	4.5 (3.5)
Overall weather conditions during the FDP VAS: 0 (very bad) to 100 (very good)	67 (25)
How demanding was the current duty day VAS: 0 (not demanding) to 100 (very demanding)	46 (18)
Effort to perform all landings VAS: 0 (extreme effort) to 100 (absolutely no effort)	73 (20)
Rating of overall workload during the FDP VAS: 0 (very low) to 100 (very high)	47 (22)

Table 10 Ratings of vigor/alertness (GV), sleepiness (SSS), and scores on the VigTrack task: VigRMS (Root Mean Square of tracking error) on days off and duty days. Means and ±SD ( ). Higher GV scores signify higher alertness levels, higher SSS scores higher sleepiness (lower alertness). Higher VigRMS scores signify poorer tracking performance. Means of reaction time of the vigilance task (VigRT) not calculated, due to missing data.

	GV	SSS	VigRMS
Day off			
Wake up	59.1 (22.9)	2.8 (1.0)	12.7 (9.2)
Bedtime	51.1 (20.0)	3.6 (1.3)	12.6 (9.2)
Duty Day			
Wake up	63.6 (23.2)	2.4 (1.1)	13.0 (8.0)
Pre-Duty	63.7 (22.2)	2.0 (1.8)	17.3 (8.4)
Halfway-Duty	73.1 (20.0)	2.0 (1.6)	12.4 (7.7)
End-Duty	68.1 (18.0)	2.3 (0.7)	13.4 (7.6)
Bedtime	48.2 (19.7)	4.0 (1.2)	16.8 (9.2)

Highest sleepiness scores and lowest vigor scores occurred at bedtime, while tracking performance was poorest at pre-duty sessions. As mentioned earlier, statistical testing of differences between the scores in de different sessions was considered not justified.

5 Discussion

The aim of the study was to assess the effects of workload and fatigue on alertness-related flight safety of pilots engaged in Dutch North Sea helicopter operations and, if necessary, to make recommendations to improve flight safety.
To achieve the aim of the study, this section provides a detailed discussion of the relevant study results.

5.1 Discussion of CHC study results

5.1.1 Pre-Duty sleep at home

Both on days off and duty days, mean GSQS scores indicated that sleep quality was good and the mean quality scores were similar to the mean scores found in a healthy Dutch population [Simons & Valk, 1998]. In a study performed in 1990 among 24 pilots of KLM Helicopters it was found that 50% had moderately severe sleep problems [Smit et al., 1990]. However in the present study the average quality of pre-duty sleep reported by the pilots was satisfactory. This contrast may be explained by the difference in duty rosters as well as FTL regulations between the North Sea helicopter operations in 1990 and the present practices, which seem to provide better opportunities for sleep during the normal circadian sleep phase (night).

On average, pilots slept 6¾ hrs before duty days, which is considered as slightly too short sleep. Shorter sleep times may have been caused by early reporting times and longer travel times to De Kooy, as both factors were significantly correlated with shorter total sleep time. In the present study, pilots with duties starting in the morning had a mean subjective total sleep time (TST) of 5:55 h (objective: 5:13 h), which is 1-1½ hour shorter than the TST of pilots starting their duties after noon. In TNO studies among 481 commercial fixed-wing pilots, it was found that alertness at top of descent (TOD) significantly impaired when the duration of pre-duty sleep was less than 7 hours [Valk et al., 2003]. There is sufficient scientific evidence that the length and quality of pre-duty sleep are important determinants of pre-flight and in-flight alertness [e.g. Carskadon & Dement, 1981; Carskadon & Dement, 1982; Rosekind et al. 1992; Pascoe et al. 1995; Valk & Simons, 1998; Dinges et al. 1997]. In the present study poor sleep quality and shorter pre-duty sleep were associated with lower vigor scores, higher sleepiness levels, and a lower level of vigilance performance.

The mean pre-duty sleep length found in the present study differs only 15 minutes with the minimum total sleep time that is considered as sufficient (viz. 7 hrs). However, we consider that a mean pre-duty TST of 6 hrs or less, as is found in pilots starting their duty in the morning, may lead to undesirable levels of alertness in some cases. Therefore, we emphasize the importance of sufficient pre-duty sleep, although in the present study sleepiness and vigilance levels never approached risk levels during the FDP.

With regard to pre-duty sleep, the findings of the present study are in complete agreement with the result of a previous study of North Sea helicopter pilots based in Aberdeen [Gander et al. 1994]. In the present study as well as in the study of Gander et al. it was found that prior to an early duty pilots went to bed earlier, but woke up earlier and consequently slept approximately one hour shorter than on their days off, or prior to an afternoon duty.

5.1.2 *Travel time from home to the airport (De Kooy)*

As has been mentioned earlier, longer pre-duty travel time was associated with shorter pre-duty sleep. Longer travel time from home to De Kooy was also significantly correlated with lower levels of vigilance performance before duty, and lower levels of vigilance and higher sleepiness levels halfway duty, at the end of duty, and at bedtime, although values never approached a risk zone during the FDP. These findings provide some evidence that vigilance and alertness are more affected as pilots live at greater distance from the airport where they have to report for duty. We surmise that this may be caused, at least partly, by shorter pre-duty sleep.

5.1.3 *Workload, fatigue, alertness, and vigilance during FDPs*

As can be expected, higher experienced workload was associated with longer duty duration, more flight hours, and a higher number of landings. Worse weather conditions were significantly correlated with more effort to perform the landings. Flying the AW139 helicopter was associated with lower scores on work demand, workload, and effort to perform the landings, compared with the S76 and EC155. It therefore appears that the AW139, which is the newest type of helicopter in the CHC fleet, may have favourable flying and/or comfort characteristics. The S61was not included in the statistical analysis due to the fact that it was only frequently flown by 2 of the 24 pilots.

Subjective data of vigor and sleepiness levels as well as objective data of vigilance performance showed high consistency. Higher workload was associated with lower vigor scores halfway-duty and at the end of duty. Highest sleepiness levels and lowest vigor scores were found at bedtime, which is a normal phenomenon, caused by normal fatiguing effects of a working day in combination with sleep pressure dictated by the circadian clock. Vigor scores were always high during the FDP and mean pre-duty, halfway-duty, and end-duty sleepiness scores were never higher than 2.3, where level 2 stands for ‘functioning at a high level, but not at peak; able to concentrate’ and level 3 signifies ‘relaxed; awake; not at full alertness; responsive’. As level 4 (‘a little foggy; not at peak; let down’) can be considered as the first level where alertness-related flight safety approaches the risk zone, it can be concluded that alertness-related safety was always maintained at a safe level during the FDPs. Frequency distributions of the scores and performance data indicate that it is unlikely that in individual cases risky sleepiness levels would have occurred.

Compared with pre-duty levels, mean alertness (vigor) and sleepiness levels were not significantly changed at halfway duty, while vigilance performance impaired with only 6.8% (tracking) and 3.8% (reaction time) compared to pre-duty levels. This small, statistically not significant, impairment of vigilance performance can be attributed to a normal process in which vigilance impairs with increasing time-on-task. This is illustrated by the findings of the present study, which show that a higher number of flight hours was associated with lower levels of vigilance performance halfway duty and at the end of duty. More flight hours were also associated with higher sleepiness levels at the end of duty, although it should be mentioned that vigilance performance or sleepiness levels never approached risk levels during a FDP.
At the end of duty vigilance was approximately at the same level as pre-duty values (tracking was even better: -6.8%; reaction time a bit slower: +1.6%). Based on results of fixed-wing studies of the European Committee on Aircrew Scheduling and Safety (ECASS), the decrease in alertness caused by time-on-task in fixed-wing operations is estimated to be approximately 10% per 2.3 flight hours [Spencer & Robertson, 2004].

In the present study, mean levels of vigor, sleepiness, and vigilance were always very favorable throughout the daily FDPs and never approached risk levels.

The levels of fatigue can be estimated using the Need for Recovery [NFR; Jansen et al., 2003] scores and the vigor and sleepiness scores, which are know to correlate highly with fatigue ratings [Samn & Perelli, 1982]. The low Need for Recovery scores at bedtime and the vigor and sleepiness scores during the FDP indicate that fatigue levels never exceeded levels that are known to negatively influence safety or health.

It can be concluded that during FDPs, pilots always maintained alertness and vigilance performance levels that are considered safe in terms of alertness-related flight safety. Compared with fixed-wing operations, alertness and vigilance in North Sea helicopter pilots is much better preserved during a FDP. This favorable observation may be caused by the facts that North Sea helicopter pilots experience less (or no) in-flight monotony and spend less duty time during the circadian sleep phase (between 00:00 and 06:00 h) compared to the fixed-wing pilots that were assessed in previous studies.

5.1.4 Summer versus Winter Operations

Summer operations showed no significant differences compared with winter operations in terms of experienced workload, how demanding work was, or how much effort it had cost. Although not statistically significant, effort to perform the landings was experienced as slightly higher during winter than in summer. At halfway-duty vigilance performance was slightly better in summer than in winter and at bedtime vigor scores were significantly higher in summer. However, summer vigor scores at bedtime were also significantly higher on days off. Lower vigor scores, indicating lower levels of energy and alertness, at bedtime in winter cannot be solely attributed to a higher workload on winter duty days, because bedtime vigor scores were also lower on days off. Otherwise, there were no significant summer-winter differences in vigor, sleepiness, and vigilance performance at wake-up, pre-duty, halfway-duty, and end-duty sessions.

Only a minor difference between summer and winter weather conditions, with slightly worse conditions in summer, was reported and this may explain that the above-mentioned (potential) differences were small to absent. Weather data of the Royal Netherlands Meteorological Institute confirm that weather conditions during the summer study period may be considered as slightly worse [KNMI, 2008].

Both on days off and on duty days, pilots slept longer and better in winter than in summer, which may be explained by a physiological human seasonality phenomenon, due to changes in photoperiod and ambient temperature [Wirz-Justice et al.,1984, Wehr, 1991; Wehr et al., 1993]. This physiological mechanism might also explain the higher bedtime vigor scores found in summer.

5.2 Discussion of Bristow study results

Because the number of pilots working in the Dutch branch of Bristow Helicopters is much smaller than the number of pilots working for CHC, it has been difficult to include a satisfying number of Bristow pilots. This was further complicated by a mismatch between availability of pilots and availability of study equipment, because of training periods and detachment abroad. Because a new type of helicopter (EC155) was introduced in the Bristow operations, all potentially participating Bristow pilots had to follow a conversion course abroad. Due to the small number of participating Bristow pilots, their data were only descriptively analysed.

Considering the Bristow data, it can be seen that the mean number of landings was considerable smaller than found in CHC operations (4.5 versus 10), which may indicate that – at least for the participating pilots – Bristow operations differed from the CHC operations. It can also be seen that weather conditions in the Bristow study period (October 2007) were reported as less favourable compared to the study periods of CHC. The vigor, sleepiness, and tracking performance scores of the five participating Bristow pilots indicate that alertness and fatigue levels during the FDP never approached risk levels concerning alertness-related flight safety.

6 Conclusions and recommendations

Conclusions and recommendations concerning CHC North Sea Operations

1 We conclude that during the present North Sea helicopter FDPs, pilots always maintained alertness and vigilance levels that are considered safe in terms of alertness-related flight safety. We also conclude that fatigue levels never exceeded levels that are known to negatively influence safety or health.
2 Pilots who had to report for an early duty had considerably shorter pre-duty sleep than others. Shorter sleep was associated with higher sleepiness levels and lower vigilance scores during the FDPs, although levels of sleepiness and vigilance never approached the risk zone associated with alertness-related safety. It is recommended to pursue an increase of the total pre-duty sleep time in pilots scheduled on morning duties. This may be done by stimulating awareness, both of pilots and management, of the importance of sufficient sleep and to try and guarantee sufficient pre-duty sleep opportunities, particularly before an early duty.
3 Pilots with longer travel time to airport De Kooy had lower levels of vigilance and higher sleepiness levels during and after the FDP. These pilots also had shorter pre-duty sleep, which may signify the causal relation with lower vigilance and higher sleepiness. Although in these cases, vigilance and sleepiness values never approached a risk zone of alertness-related safety, it is recommended to stimulate opportunities for shorter travel time from home to De Kooy, whenever socially acceptable and possible.
4 Differences between summer and winter operations were minimal in terms of workload, vigilance, and sleepiness scores. Weather conditions were experienced as only slightly worse during the summer, which may explain the lack of differences between both seasons. Both on days off and on duty days, pilots slept longer and better in winter than in summer and had lower bedtime vigor scores in winter. Both phenomena are considered as normal human seasonality effects.
5 Flying the AW139 helicopter was associated with lower scores on work demand, workload, and effort to perform the landings, compared with the S76 and EC155.

Conclusions concerning Bristow North Sea Operations

Vigor, sleepiness, and tracking performance scores of the five participating Bristow pilots indicate that alertness and fatigue levels during the FDP never approached risk levels concerning alertness-related flight safety. The small amount of data did not allow for more detailed conclusions, such as those described in the CHC part of this section.

Overall conclusion of the project (phase 1 and 2)

Taking into account the conclusions of the report of phase 1 (TNO DV 2008 C027) and above-mentioned conclusions, we conclude that flight safety in North Sea helicopter operations may be primarily affected by other factors than low alertness or high fatigue levels. In this context, principal factors are turbulence around offshore platforms, unfavourable weather conditions, distraction by calculating load/fuel or filling in paper work, and poor helideck lighting. Additional unfavourable factors may be wearing exposure suits, light conditions, cockpit misting, and the proficiency level of the Helicopter Landing Officer. Cabin environmental conditions, such as vibrations, ventilation and temperature, and unfavourable body positions may contribute to health problems.
In that context, it would be useful to assess whether the introduction of new types of helicopters, such as the EC 155 and AW 139, will reduce health problems that were always considered to be specific for helicopter pilots (e.g. back pain).

7 References

Ancoli-Israel, S.; Cole, R.; Alessi, C.; Chambers, M.; Moorcroft, M. & Pollak, C.P. (2003), The role of actigraphy in the study of sleep and circadian rhythms, *Sleep, 26(3)*:342-392.

Carskadon, M.A. & Dement, W.C. (1981), Cumulative effects of sleep restriction on daytime sleepiness, *Psychophysiology; 18*:107-113.

Carskadon, M.A. & Dement, W.C. (1982), Nocturnal determinants of daytime sleepiness, *Sleep; 5*:S73-S81.

Dinges, D.F.; Pack, F.; Williams, K.; Gillen, K.A.; Powell, J.W.; Ott, G.E.; Aptowicz, C. & Pack, A.I. (1997), Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4-5 hours per night, *Sleep; 20(4)*:267-7.

Gander, P.H.; Barnes, R.M.; Gregory, K.B.; Conell, L.J.; Miller, D.L. & Curtis Graeber, R. (1994), *Crew Factors in Flight Operations VI: Psychophysiological responses to Helicopter Operations*. NASA Technical Memorandum 108838. July 1994, NASA Ames Research Center, Moffett Field, CA.

Hoddes, E.; Zarcone, V.; Smythe, H.; Philips, R. & Dement, W.C. (1973), Quantification of sleepiness: a new approach, *Psychophysiology, 10*: 431-36.

Jansen, N.W.H.; Kant, I.J.; Amelsvoort, L.G.P.M. van; Nijhuis, F.J.N. & Brandt, P.A. van den (2003), Need from recovery from work: evaluating short-term effects of working hours, patterns and schedules, *Ergonomics 46(7)*:664-680.

KNMI (2008), *http://www.knmi.nl/klimatologie/maand_en_seizoensoverzichten/index.html*

Meijman, T.F.; Vries, G.M. de; Vries-Griever, A.H.G. de & Kampman, R. (1987), *The evaluation of the Groningen Sleep Quality Scale. Department for experimental and occupational psychology*, Biological Centre, University of Groningen 1987, The Netherlands.

Monk, T.H. (1989), A visual analogue scale technique to measure global vigor and affect, *Psychiatry Research; 27*:89-99.

Pascoe, P.A.; Johnson, M.K.; Robertson, K.A. & Spencer, M.B. (1995), *Sleep in rest facilities on board aircraft: Field studies*, DRA Report No DRA/CHS/A&N/CR/95/002, March 1995.

Rosekind, M.R.; Graeber, R.C.; Dinges, D.F.; Conell, L.J.; Rountree, M.S.; Spinweber, C.L. & Gillen, K.A. (1992), *Crew Factors in Flight Operations: IX. Effects of preplanned cockpit rest on crew performance and alertness in long-haul operations*. NASA Technical Memorandum 103884, NASA-Ames Research Center, Moffett Field, CA.

Sadeh, A.; Alster, J.; Urbach, D. & Lavie, P. (1989),
Actigraphically based automatic bedtime sleep-wake scoring: validity and clinical applications,
J. Ambulatory Monitoring. 2(3):209-16.

Samn, S.W. & Perelli, L.P. (1982),
Estimating aircrew fatigue: A technique with implications to airlift operations,
Brooks AFB, TX: USAF School of Aerospace Medicine. Technical Report No. SAM-TR-82-21, 1982.

Simons, M.; Valk, P.J.L.; Ree, J.J.D. de; Veldhuijzen van Zanten, O.B.A. & D'Huyvetter, K. (1994),
Quantity and quality of onboard and layover sleep: effects on crew performance and alertness. Report RD-31-94,
Netherlands Aerospace Medical Centre, Soesterberg.

Simons, M. & Valk, P.J.L. (1997),
Effects of a Controlled Rest on the Flight Deck on Crew Performance and Alertness. Report: NLRGC 1997-B3,
Netherlands Aerospace Medical Centre, Soesterberg.

Simons, M. & Valk, P.J.L. (1998),
Early starts: effects on sleep, alertness and vigilance. AGARD-CP-599,
NATO-AGARD, Neuilly-sur-Seine, France. p. 6/1-6/5.

Simons, R.; Koerhuis, C.L.; Valk, P.J.L. & Oord, M.H.A.H. van den (2006),
Usefulness of Temazepam and Zaleplon to Induce Afternoon Sleep,
Military Medicine; 171:998-1001.

Simons, M. (2008),
Effecten van werkdruk en vermoeidheid op vliegveiligheidsaspecten en gezondheid van vliegers in Noordzee helicopteroperaties – literatuurstudie. Rapport TNO-DV 2008 C027,
TNO Defensie en Veiligheid, Soesterberg.

Smit, J.; Bruggink, B.; Meer, O van der & Valk, P.J.L. (1990),
Effects of workload on off-shore helicopter pilots – a field study. Report 90-01-RLD,
Netherlands Aerospace Medical Centre, Soesterberg, The Netherlands.

Spencer, M. & Robertson, K. (2004),
Aircrew alertness on the Singapore–Los Angeles route: final report. Qinetiq Report no QINETIQ/KI/CHS/CR050022/1.0. December 2004.

Valk, P.J.L.; Simons, M.; Struyvenberg, P.A.A.; Kruit, J. &
Van Berge Henegouwen, M. (1997),
Effects of a single dose of loratadine on flying ability under conditions of simulated cabin pressure,
American Journal of Rhinology, 11(1), 27-33.

Valk, P.J.L. & Simons, M. (1998),
Pros and cons of strategic napping on long haul flights. AGARD-CP-599,
NATO-AGARD, Neuilly-sur-Seine, France (pp. 5/1-5/5).

Valk, P.J.L.; Roon, D.B. van & Simons, M. (2001),
Effects of an alcohol hangover on pilot's performance and alertness,
Nederlands Militair Geneeskundig Tijdschrift; 54:4-10.

Valk, P.J.L.; Simons, M.; Goei, J.H. & Hijum, S.M. van (2003),
Evaluation of the Fit-to-Fly Checklist on Long Haul and Short Haul Flights. Proceedings 15th EASS – FSF & ERA Conference 'Change: A safety Challenge', March 2003, Geneva, Switzerland,
Flight Safety Foundation, Alexandria, Virginia USA. pp.9-16.

Valk, P.J.L.; Roon, D.B. van; Simons, M. & Rikken, G. (2004),
Desloratadine shows no effect on performance during 6h at 8,000ft simulated cabin altitude,
Aviat Space Environ Med;75:433-438.

Veldhoven, M. van & Broersen, S. (2003),
Measurement quality and validity of the 'need for recovery scale',
Occup Environ Med; 60(Suppl 1): i3–i9.

Wehr, T.A. (1991),
The durations of human melatonin secretion and sleep respond to changes in daylength (photoperiod),
J Clin Endocrinol Metab, 73:1276-1280.

Wehr, T.A.; Moul, D.E.; Barbato, G.; Giesen, H.A.; Seidel, J.A.; Barker, C. & Bender, C. (1993),
Conservation of photoperiod-responsive mechanisms in humans,
Am J Physiol, 265:R846-R857.

Wirz-Justice, A.; Wever, R.A. & Aschoff, J. (1984),
Seasonality in freerunning circadian rhythms in man,
Naturwissenschaften, 71:316-319.

8 Signature

Soesterberg, December 2008

TNO Defence, Security and Safety

J.P. Dezaire, MSc
Head of department

M. Simons, MD
Author

A Instruction

Dear Pilots,

TNO Aerospace Medicine is asking for your participation in a study among pilots of CHC and Bristow. The study is commissioned by the Dutch Aeronautical Inspection (insp. Bert Vos and Wandert Brandsen).

The aim of this study is:
To assess the effects of current flight and duty time limitations and rest schemes, working environment, and fatigue on flight safety aspects and health of pilots engaged in North Sea Helicopter OPS. If necessary, recommendations will be made to enhance flight safety and health of pilots.

Method
We conduct a literature study (what is already known?) and a field study. In the field study we will assess your fatigue and alertness levels, and sleep quality variables during duty periods and days off. There will be two 14-days assessment periods (Summer and Winter).

We will equip you with an Actiwatch (right photograph) and a PDA (left photograph), which you will have to use each day of the 14-days study period. The Actiwatch is used to assess sleep and has the size of a watch. It has to be worn around the wrist during the day and night. It does not cause any discomfort and does not interfere with normal activities. The PDA is used to answer questions about fatigue, sleep, and operational characteristics and to perform an 5-min alertness test. We will instruct you how to use both devices and you will be trained on the alertness test. We will collect the devices after each 14-day test period and download your results. Participation is on a voluntary basis and you can withdraw at any moment, without reason given.
Your data will be processed anonymously. All personal information is subject to the duty of professional confidentiality of the project leader and physician Ries Simons.

What do we ask from you?
– Once, at the start of the test period: to be available for 45-60 min instruction and training.
– To wear the Actiwatch each day during your sleep period (days off + duty period).
– To answer the questions on the PDA and to perform the alertness test: each work day max. 5 test sessions 1) after getting up, 2) before start of duty, 3) at a quiet moment halfway your duty, 4) after the end of duty, and 5) at bedtime (exact timing and number of sessions will depend on your duty scheme). Each test session (questions + alertness test) will take no longer than 10 minutes. On your days off, we ask you to perform 2 test sessions: after getting up and at bedtime.
  It is important to perform the test sessions in a quiet environment, where you can fully concentrate on the test.
– We ask you to participate in both 14-days periods. We expect differences in workload, in particular caused by differences in light conditions between the two study periods.

We are aware that we ask much of your efforts and motivation. But we really need you, because your on- the-job experiences are indispensable to optimize health and safety in the sector.

Since 1988, we run the ‘Aircrew Fatigue Countermeasures Programme’ commissioned by the Dutch Aeronautical Inspection. In 2002 our research department has been transferred from the Aeromedical Institute to TNO, both seated in Soesterberg. In the fixed wing sector, we conducted studies of fatigue, alertness, sleep, and health in short-haul, long-haul, and ultra-long-haul operations (involving over 500 pilots). In these studies we used the same methods as we are using in the current study. We are actively collaborating in the international consortium ECASS (European Committee on Aircrew Scheduling and Safety).

We will present the results of this study to you as well as to the Ministry of Transport (Aviation Dept.) The report will be available to you. The results will not be traceable to individual pilots.

We are looking forward to meet you soon for instruction and distribution of the equipment.
Please do not hesitate to contact us in case you have any questions.

Ries Simons ph: +31 (0)346-356485 e-mail: ries.simons@tno.nl
Pierre Valk ph: +31 (0)346-356211 e-mail: pierre.valk@tno.nl

After the instruction, pilots received the following written instruction:

Actiwatch: every day
Put Actiwatch on when you go to bed
Press MARKER button at “lights out”
Press MARKER button when awakening at end of sleep period
Take Actiwatch off before taking a shower

PDA: every day-off (select day off in menu)
1. Test session 30-60 min after getting up
2. Test session at bedtime

PDA: every reserve day (select reserve day in menu)
1. Test session 30-60 min after getting up
2. Test session at bedtime

PDA: every working day (select working day in menu)
1. Test session 30-60 min after getting up
2. Test session before start duty (only if more than 2 hours between session 1 and 2)
3. Test session ± halfway your duty (take a quiet moment!)
4. Test session at end of duty
5. Test session at bedtime

Please use the charger to load the batteries every night

B Global Vigor and Affect Scale [GVA; Monk, 1989]

Global Vigor and Affect Scale [GVA; Monk, 1989]
How alert do you feel? very little (0)--(100) very much
How sad do you feel? very little (0)--(100) very much
How tense do you feel? very little (0)--(100) very much
How much of an effort is it to do anything? very little (0)--(100) very much
How happy do you feel? very little (0)--(100) very much
How weary do you feel? very little (0)--(100) very much
How calm do you feel? very little (0)--(100) very much
How sleepy do you feel? very little (0)--(100) very much

The formulas used are:

GV = [(alert) + 300 – (sleepy) – (effort) – (weary)] / 4
GA = [(happy) + (calm) + 200 – (sad) – (tense)] / 4

Each formula yields a value between 0 and 100 which is rounded to the nearest whole number.

C Stanford Sleepiness Scale [SSS; Hoddes et al., 1973]

Stanford Sleepiness Scale [SSS; Hoddes et al., 1973]	
1	Feeling active and vital; alert; wide awake.
2	Functioning at a high level, but not at peak; able to concentrate.
3	Relaxed; awake; not at full alertness; responsive.
4	A little foggy; not at peak; let down.
5	Foggy; beginning to lose interest in remaining awake; slowed down.
6	Sleepy; prefer to be lying down; fighting sleep; woozy (NL: suf).
7	Almost in reverie; sleep onset soon; losing struggle to remain awake.

D Need for Recovery Scale [NFR; Jansen et al. 2003]

Need for Recovery Scale (NFR; Jansen et al. 2003)			
		yes	no
1	I was hard for me to relax after work today.		
2	At the end of the duty today, I was really tired.		
3	Today I felt rather exhausted after work.		
4	After dinner, I felt rather fit.		
5	I did not come to ease this evening.		
6	It was hard for me to concentrate in the free time after my work.		
7	Directly after returning from work today, it was hard for me to pay attention to other people.		
8	It took me more than one hour to recover from work today.		
9	After returning home from work, I needed to be left alone for a while.		
10	Today I felt too fatigued after work to do something in the free time.		
11	My performance on the last part of my duty was not optimal due to fatigue.		

The answer ‘yes’ signals unfavourable situations and scores 1, except for item 4, where ‘no’ signals an unfavourable situation and scores 1. Total score is divided by the number of items x 100. Higher scores represent higher need for recovery.

Distribution list

The following agencies/people will receive a complete copy of the report.

5 ex.	Inspectie Verkeer en Waterstaat – Luchtvaart W. Brandsen
2 ex.	TNO Defensie en Veiligheid, vestiging Soesterberg, Archief
3 ex.	TNO Defensie en Veiligheid, vestiging Soesterberg, M. Simons, Arts drs. E.S. Wilschut drs. P.J.L. Valk