INITIAL EVIDENCE-BASED ANALYSIS OF RISK LEVELS PER PHASES OF FLIGHT, RECENCY REQUIREMENTS AND VISUAL APPROACH DESIGN IN NIGHTTIME OFFSHORE HELICOPTER OPERATIONS

Mr. Felipe Nascimento Dr. Arnab Majumdar Prof. Washington Ochieng Dr. Wolfgang Schuster
Imperial College London
The Lloyd's Register Foundation Transport Risk Management Centre
Centre for Transport Studies
United Kingdom

Abstract

The analysis of risks per phases of flight is fundamental for safe nighttime offshore helicopter operations. However, the poor quality of the safety data currently available limits the identification of critical issues on a per-phase-of-flight basis. To redress this problem, this paper develops a customised taxonomy of phases of nighttime offshore helicopter flights and uses it as the basis for a questionnaire survey on the phase-specific risk levels experienced by pilots. Additionally, the critical issues of flying recency requirements and visual approach design preferences are investigated. With the responses obtained from pilots located in seven countries, extensive statistical hypothesis testing shows that the phases involving visual scan techniques at high speed regimes are problematic, especially the visual segment of instrument approaches. Furthermore, the between-night-flights time gaps required for assured recency were found considerably shorter than currently standardised across the industry. A number of important implications result and should form the basis for future safety interventions.

1. INTRODUCTION

Phases of flight reflect comparable aerodynamic loads, environmental hazards and pilot activities. As such, phases of flight receive special consideration during the investigation of safety occurrences [1-6]. However, the phases of helicopter flight are still poorly discriminated because the taxonomies in use are ill-adapted from fixed-wing operations [7-9]. This impacts on the quality of the safety data collected on the basis of phases of flight [6, 10-11] and prevents the identification of specific problem areas in helicopter operations.

Using clustered phases of flight is the typical solution to deal with data collected on the basis of poorly discriminated phases. This is widely applied in offshore helicopter safety analysis, leading to a concentration of the associated interventions on the clustered approach-and-landing phase. Recent interventions have aimed to support operations in degraded visual environment, especially the nighttime [12], which is associated with significantly higher accident rates and increased odds of pilot-related and fatal accidents [6]. The nighttime is also associated with rapid decay of piloting skills and various procedural inconsistencies, especially in the final visual portion of the approach [1, 13-14].

However, the current focus of safety interventions on the approach-and-landing clustered phase stems from safety concerns identified primarily in the North Sea in the 80s and 90s [12], and an updated assessment of the risks across more well-discriminated phases is now required globally [6]. This cannot be undertaken on the basis of the analysis of accident and incident reports due to low data quality. Similarly, ideal training recency requirements cannot be derived from the data available, which equally do not support the identification of any visual approach procedural design which might be preferred by pilots for being safer.

To address the shortcomings mentioned this paper aims to assess the risk levels associated with the phases of nighttime offshore helicopter flights through a questionnaire survey applied to helicopter pilots. It also aims to investigate the intervening time which assure that the piloting skills required do not decay to an alarming level between night flights and any preferences for visual approach procedure design.

A critique of the limitations of the taxonomies and previous analyses undertaken on the basis of phases of flight is required.

2. TAXOMIES AND ANALYSES ON THE BASIS OF PHASES OF FLIGHT

The logic underpinning the development of current taxonomies of phases of flight is usually unclear. Whilst some appear to describe the spatial position of the aircraft, others focus on the tasks of the aircrew or events detectable by Flight Data Monitoring (FDM) devices [4, 7, 15-16]. Most taxonomies prescribe excessively broad phases of flight and, hence, can be criticised for a lack of explanatory power [1, 6, 10-11, 17-18]. Customised taxonomies redress this problem; however, these are few and usually not in the helicopter domain [5]. From all said, comparing the results of studies based on phases of flight is currently both difficult and potentially misleading.

The studies based on the phases of flight display yet another shortcoming: the limited exploration of the statistical relationships between the variables relevant to safety, especially pilot demographics, and the phases of flight [6, 10]. As a consequence, the generalisability of such studies is questionable.

The methodology described in the next section overcomes these problems by assessing the perceived risk levels associated with the phases of nighttime offshore helicopter flights, exploring the statistical relationships with demographic variables of the respondents and investigating nighttime flying recency and visual approach design preferences. These tasks were undertaken through a questionnaire survey.

4. METHODOLOGY

The questionnaire comprises of subjective risk level ratings for the different phases of a typical nighttime flight, for which a customised taxonomy was developed. Each pilot rated all phases and additionally replied to a set of demographic questions and two direct questions covering the issues of time gap required for skill decay and visual approach procedure design. The questions were piloted and a sampling strategy was devised to ensure the generalisability of the results. Details of each of such steps are reported in the sections below.

4.1 Taxonomy of phases of nighttime flights

The phases of offshore helicopter flights proposed in [16] were improved and expanded with the process charting technique [5] and the hierarchical task analysis of [13]. This led to a breakdown of the approach, takeoff and go-around manoeuvres into visual and instrument segments.

The resulting phases were then clustered into two super-ordinate levels (labelled levels 2 and 3) according to the clustering strategies in [6, 10-11]. To account for the specific characteristics of nighttime operations, categories that reflected the helicopter's kinetic states (high versus low speed, hence fast versus slow variation of flight parameters, respectively) and pilot's visual sampling strategies (visual versus instrument scans) were also introduced. This enabled to identify the effects of human factors shortcomings more accurately. The resulting taxonomy is shown in Table 1.

Table 1 – Taxonomy of phases of flight

Phase of flight			Scan technique (St)	Kinetic state (Ks)	Combined St - Ks
Level 1	Level 2	Level 3			
Parked, rotors/engines not running	Parked	Ground manoeuvres	Visual (V)	Low (L)	VL
Parked, rotors/engines running	Parked	Ground manoeuvres	Visual (V)	Low (L)	VL
Hovering before positioning/taxiing for takeoff	Taxiing	Ground manoeuvres	Visual (V)	Low (L)	VL
Positioning/taxiing before takeoff	Taxiing	Ground manoeuvres	Visual (V)	Low (L)	VL
Visual segment of takeoff/initial climb	Takeoff	Departure	Visual (V)	High (H)	VH
Instrument climb	Takeoff	Departure	Instrument (I)	High (H)	IH
Instrument cruise	Cruise	Cruise	Instrument (I)	High (H)	IH
Instrument descent	Approach	Arrival	Instrument (I)	High (H)	IH
Instrument approach	Approach	Arrival	Instrument (I)	High (H)	IH
Visual segment of the approach	Approach	Arrival	Visual (V)	High (H)	VH
Positioning/taxiing after the approach	Taxiing	Ground manoeuvres	Visual (V)	Low (L)	VL
Hovering after positioning/taxiing after approach	Taxiing	Ground manoeuvres	Visual (V)	Low (L)	VL
Landing	Landing	Arrival	Visual (V)	Low (L)	VL
Go around in the instrument segment of the approach	Approach	Arrival	Instrument (I)	High (H)	IH
Go around in the visual segment of the approach	Approach	Arrival	Visual (V)	High (H)	VH

4.2 Questionnaire design

The pilots were instructed to answer the questions according to their experience with perfectly serviceable aircraft. This retained the focus on the risks related to human factors shortcomings.

The demographic variables were collected to enable the analysis of risk ratings per specific groups of pilots and control for potential biases associated with surveys. The demographics were drawn from the literature and are specified in the results and discussion section [10].

Eleven-point rating scales [19] were semantically-anchored [20-21] to assess the subjective risk levels of each level 1 phase as follows: 0 meant 'no risk at all', 5 was the mid-point (i.e., 'somewhat risky') and 10 meant 'extremely risky'. Based on [22], risk was defined as 'the likelihood that an adverse event of damaging consequence will happen'. The level 1 phases were presented to participants together with the rating scales so that the pilots could rank the phases relatively to each other as much as in absolute terms.

The direct question addressing pilot recency asked pilots to state in how many days after a nighttime flight they felt that a second nighttime flight would be needed so that they would still feel current. This enabled to challenge the standard of the International Association of Oil and Gas Producers (OGP), whereby pilots are required to perform three nighttime approaches and takeoffs within a period of 90 days [23]. This is very important since there is evidence that the current standard is insufficient to ensure pilot recency [1, 13-14, 24].

The second direct question asked the pilots' preferences for visual approach procedure design between two options: levelled or sloping visual segments, typical of airborne radar approach (ARA) procedures and spatial gates, respectively [13-14].

The phases of flight, scales of risk, demographic, and direct questions were piloted by subject matter experts (SMEs), mostly senior offshore helicopter pilots involved in training and safety.

4.3 Pilot sampling

Because the nighttime is a high risk factor for offshore helicopter operations worldwide [6], world coverage was pursued using the areas outlined in [1, 6] as the basis for a regional sampling strategy. Group 1 was formed by mid-latitude areas of typically benign weather conditions and mostly covered by mid-twin turbine helicopters (e.g., Brazil and Spain). Group 2 was formed by Alaska and Canada, where the weather conditions are usually adverse. Group 3 was formed by the countries of the Commonwealth of Independent States (CIS), which differ from the previous group due to potentially different airworthiness requirements [25]. Group 4 was formed by the Gulf of Mexico, dominated by single engine helicopters, and Group 5 by the North Sea, where regulations are renowned for stringency.

As far as possible, the major helicopter companies, aviation regulators, accident investigation boards, oil and gas companies and safety groups [e.g., 26] in each region were contacted, with the aim to reach two groups of pilots: senior and flight line-only pilots. The former were flight instructors, aviation safety officer, operations and management post-holders, whose extensive experience was anticipated to synthesise that of many pilots [27]. The flight line-only pilot group was formed by line pilots without an administrative job. These were used to control for potential professional bias of the previous group.

Contact was established via electronic mail addressed to the persons in charge of the safety, operations and management departments of the targeted companies. In four countries mentioned in the results and discussion section, the questionnaire was primarily applied by the main author in individual and confidential face-to-face sessions. Any potential bias introduced by face-to-face contact was deemed irrelevant in light of the simplicity of the questionnaire [28]. All participants in the survey were volunteers.

The required sample sizes were estimated as in [28] using three variables: number of nighttime deck landings, risk ratings and membership to the 'senior pilot' group. The range of the number of night deck landings was assumed from the minimum of three required by the OGP for nighttime recency [23] to a maximum of 1080, which considers *inter alia* a pilot service life of 40 years. The range of risk ratings was determined from the 0-10 scales used. The desired level of confidence was set at 1.96, corresponding to 95% confidence interval. The arbitrary error accepted was 50 night deck landings and 0.5 risk rating. Finally 10% was accepted as a reasonable expression of the true proportion of senior pilots in the population of pilots. Based on these assumptions the following per group sample sizes were calculated: i) 50 based on number of night deck landings; ii) 51 based on risk ratings; iii) 35 based on seniority membership. Finally, responses from at least 51 senior and 51 flight-line only pilots were pursued.

4.4 Statistical treatment

Because the risk ratings of each phase were attributed by comparison to one another, the ratings were treated as continuous data [29]. Using statistical procedures that fitted the characteristics and distributions the data, the following hypotheses were initially tested:

- $H0_1$: the distributions of risk ratings across phases of flight (level 1) are the same;
- $H0_2$: the distributions of risk ratings across phases of flight (level 2) are the same;
- $H0_3$: the distributions of risk ratings across phases of flight (level 3) are the same;
- $H0_4$: the distributions of risk ratings across combined scan technique – kinetic state (St-Ks) categories are the same.

These hypotheses highlighted which were the most critical phases of flight. Furthermore, the tests of $H0_2$ and $H0_3$ indicated whether clustering flight phases produced misleading results.

Based on the results of the previous tests, the following hypotheses were subsequently tested:

- $H0_5$: the distributions of risk ratings across the level 1 phases forming the most critical level 3 phase (identified from the test of $H0_3$ above) are the same;
- $H0_6$: the distributions of risk ratings across the level 1 phases forming the most critical St-Ks category (identified from the test of $H0_4$ above) are the same.

Such tests challenged the results of the test of $H0_1$ and enabled the selection of the most critical phase for the subsequent demographic analysis that aimed to identify regions or groups of pilots particularly vulnerable to the most critical phases of flight. The tests were undertaken by analysing the distributions of risk ratings per demographic variable collected.

Lastly, the time gaps for recency decay and the preferred visual approach procedure design were analysed against the demographic variables. The hypotheses tested and the statistical tests applied are introduced in the results and discussion section. The hypotheses were two-tailed and significance was established at $p \leq .05$.

5. RESULTS AND DISCUSSION

5.1 Descriptive statistics

Sixty-one questionnaires were returned, of which 40 were obtained from face-to-face sessions in Norway, Holland, Brazil and Spain. Respondents were based in Groups 1 (i.e., mid-latitude countries, N=15) and 5 (i.e., North Sea, N=46).

Most pilots were male (N=58), with only three female pilots participating in the survey. There were 46 captains and 13 first officers. Fifty respondents held air transport pilot's licences (ATPL) and 9 held commercial pilot's licences (CPL); 23 were 'flight line-only pilots' and 35 respondents were 'senior pilots'. The former corresponds to an 81.3% confidence interval for both the number of night deck landings and risk ratings, which was deemed acceptable for statistical analysis. Furthermore, the sample of flight line-only roughly matched the size of the senior pilot group and, therefore, was adequate for statistical comparisons [28]. The sums that did not reach the total of 61 stem from slightly incomplete questionnaires. The main demographic characteristics of the pilots sampled is summarised in Table 2.

Table 2 – Main demographic characteristics of helicopter pilots

	Mean	Standard deviation (SD)
Age, years	43	9
Experience flying helicopters, hours total	5905	3483
Experience flying helicopters, hours night	3179	2917
Experience flying helicopters, hours IFR	887	829
Number of night deck landings	1926	3868
Length of time working as offshore helicopter pilot, years	10	9

The helicopter types flown were categorised according to transport capacity [6] (i.e., heavy twin turbine, HT, N=20; and mid twin turbine, MT, N=37) and cockpit generation [1, 11] (i.e., glass cockpit, N=47; and analogic/mixed cockpit, N=10).

The medians of the ratings assigned to the flight phases can be seen in Table 3. Medians are meaningful with rank-based non-parametric tests.

Regarding the direct questions, the mean time gap for recency decay reported was 28 days (SD=56.9) and, in response to the second direct question, 19 and 30 pilots preferred respectively the levelled and sloping visual approach segments.

5.2 Inferential statistics

The normality of all the distributions described henceforth were tested according to the sample sizes involved using either the Shapiro-Wilk's or Kolmogorov-Smirnov tests. Homogeneity of variance was checked by Levene's test [30].

5.2.1 Tests of $H0_{1\text{-}4}$

In all cases at least one assumption of parametric data was violated and, therefore, non-parametric techniques were required. Related Samples Friedman's Two-Way Analysis of Variance by Ranks were used to account for the fact that each respondent scored every phase of flight proposed [30]. The results of the tests of$H0_{1\text{-}4}$ were significant and hence the hypotheses were rejected.

Table 3 – Medians and standard deviations (SD) of ratings attributed to each phase of flight (highest medians in under script font)

Level 1	Median	Level 2	Median	Level 3	Median	Combined St - Ks[1]	Median
Parked, rotors/engines not running	0.0	Parked	1.3	Ground manoeuvres	3.5	Visual scan - low energy (VL)	3.0
Parked, rotors/engines running	2.0	Taxiing	4.0	Departure	4.5	Visual scan - high energy (VH)	6.0
Hovering before positioning/taxiing for takeoff	4.0	Takeoff	4.5	Cruise	3.0	Instrument scan - high energy (IH)	3.0
Positioning/taxiing before takeoff	3.0	Cruise	2.0	Arrival	4.7		
Visual segment of takeoff/initial climb	5.0	Approach	4.6				
Instrument climb	4.0	Landing	5.0				
Instrument cruise	3.0						
Instrument descent	3.0						
Instrument approach	4.0						
Visual segment of the approach	6.5						
Positioning/taxiing after the approach	5.0						
Hovering after positioning/taxiing after approach	4.0						
Landing	5.0						
Go around in the instrument segment of the approach	5.0						
Go around in the visual segment of the approach	6.0						

5.2.2 Post hoc analysis of $H0_{1-4}$

The Related Samples Wilcoxon Signed Rank Test with Bonferroni correction was applied. The Bonferroni correction avoided falsely rejecting the null hypothesis by using a more stringent rejection criterion given by the initial significance level divided by the number of pairs under analysis. In each such test, the category with highest median was chosen as the control group [6, 10] to enable an assessment against the group with potentially the highest risk ratings. Therefore, the relevant results are the statistically insignificant results because they indicate phases in which risk ratings are comparable to the worst of the phases.

The post-hoc tests of $H0_1$ led to 14 pairs of phases being formed for comparison. The only ratings that failed to achieve significance, and hence were comparable to those attributed to the 'visual segment of the approach' (i.e., control group), were those of the 'visual segment of the takeoff', 'go around in the instrument segment of the approach' and 'go around in the visual segment of the approach'.

The post hoc tests of $H0_2$ generated five pairs of phases for comparison. The only ratings that failed to achieve significance, and hence were comparable those attributed to the 'landing' level 2 phase (i.e., control group), were those of the 'taxiing', 'takeoff' and 'approach' phases. Hence, the clustering of flight phases as level 2 phases appeared to have generated two misleading results: firstly, 'landing' appeared as the most critical level 2 phase (i.e., highest median). Secondly, 'taxiing' figured as a level 2 phase of concern even though no level 1 categories within it achieved significance during the previous tests. This endorses the view that care should be taken when interpreting studies which use clustered phases of flight, including official studies [6, 10, 18].

The post hoc tests of $H0_3$ were undertaken as three paired-tests, during which only the 'departure' phase failed to show significant differences in pilot ratings ($p>.05/3$) in comparison to the 'arrival' level 3 phase (i.e., control group). Hence, the risk levels assigned were comparable between the 'arrival' and 'departure' level 3 phases.

The post hoc tests of $H0_4$ compared the ratings of the VL and IH phases against those of the VH phase (i.e., the control group). Both tests were significant, and therefore the risk ratings were significantly higher for the VH phases.

5.2.3 Tests of $H0_{5-6}$

The tests of $H0_5$ and $H0_6$ tested the distributions of risk ratings across the level 1 phases within the 'arrival' (level 3) and 'VH' phases, respectively. The statistical tests mentioned above were applied because of similar data characteristics and sample sizes.

The test of $H0_5$ was significant. The post hoc tests revealed that high ratings in the 'visual segment of the approach' (i.e., control group) were only comparable to those of the 'go around in the visual segment of the approach' (i.e., $p>.05/5$). This challenges the post hoc tests of $H0_1$, which indicated that the 'go around in the instrument segment of the approach' also belonged to this population. However, the latter's inclusion was likely a type II error during the post hoc tests of $H0_1$, stemming from the excessively stringent rejection criterion caused by the application of the Bonferroni correction to 14 pairs of phases.

Considering that the test of $H0_6$ failed to produce significant results, and hence the risk ratings assigned to all VH phases were similar, the phases rated highest for risk were those flown by reference to external visual cues at high kinetic energy states (VH, i.e., the visual segments of takeoff, approach and go around). Considering these results in light of the worldwide concentration of offshore helicopter pilot-related accidents in the arrival segment of flight [6, 10] and the numerous international efforts to redress this problem, the 'visual segment of the instrument approach' can be considered to be the single most critical phase of nighttime offshore helicopter flights. Therefore, the statistical analysis that follows focuses on both this phase and the VH category.

5.2.4 Risk ratings of the visual approach and VH phases per demographic variables

Non-parametric tests were needed during the demographic analysis due to the characteristics and distributions of the data. The Independent Samples Mann-Whitney U Test was used to compare the distributions of risk ratings across categories of the following nominal variables: region, helicopter transport capacity, helicopter cockpit generation, pilot's licence, pilot's gender, pilot's rank, pilot's category of seniority (i.e., senior pilot or flight line-only pilot) and pilot's preferred visual approach design. The Kendall's Tau b non-parametric correlations were used against the following continuous variables: pilot's age, pilot's time as an offshore helicopter pilot, pilot's total flying hours in helicopters, pilot's IFR flying hours in helicopters, pilot's night flying hours in helicopters, pilot's number of night deck landings, pilot's suggested time gap for recency decay.

Significance was only obtained for the distributions of VH risk ratings per categories of helicopter transport capacity and cockpit generation. These occurred as the risk ratings were higher for mid twin-turbine and glass cockpit helicopters, respectively. Because the results were insignificant for the 'visual segment of the approach', the significances stem from the visual segments of either the takeoff or go around.

5.2.5 Time gap for recency decay per demographic variables

Given the similar data characteristics and sample sizes involved, the statistical tests mentioned above were applied to test the distributions of time gap across the demographic variables of the previous sections, except for the time gap for recency decay.

The results were significant for helicopter transport capacity, helicopter cockpit generation, category of pilot seniority and number of night deck landings. Shorter time gaps for recency decay were reported by heavy twin-turbine (HT) helicopter pilots. This was unexpected given that the pilots of mid-twin turbine (MT) helicopters rated risk higher for the VH phases and, therefore, it would have been natural that shorter time gaps would have been suggested by MT pilots. This contradiction might indicate that the risk experienced did not relate to the VH phases, that MT pilots had low awareness of practicing needs or potentially displayed some 'locus of control' attitude [1, 11, 31-32] for example. It is also possible that such helicopters instigate complacent behaviour, which might in turn result in unjustifiably over-stretched envelopes [10]. The reasons for this discrepancy between risk rating and time gap for recency decay require further investigation.

The significance with respect to cockpit generation occurred as pilots of glass cockpit helicopters mentioned shorter time gaps for recency decay. This indicates that flying glass cockpits is not a safety panacea and might be more cognitively demanding than assumed in current training requirements [33].

The significance regarding category of pilot seniority was associated with senior pilots (as opposed to flight line-only pilots) mentioning larger time gaps for recency decay (i.e., senior pilots' mean = 52.1 days, SD = 75.9; flight line-only pilots' mean = 6.2 days, SD = 7.2). This could stem from a sense that pilots with a larger overall experience can afford practicing night flights less frequently and still feel current, or might reflect a professional bias of senior pilots. By being involved with regulatory authorities and the OGP, senior pilots might have displayed a tendency to comply with the 90-day nighttime recency standard currently applied in most countries [13-14, 23].

The significance on the number of night deck landings resulted from the positive correlation between the number of night deck landings and the time gap for recency decay. The explanations over the significance of pilot seniority are equally plausible since senior pilots will naturally have accrued more nighttime deck landings over their careers.

5.3.6 Preferred visual approach design per demographic variables

Given the data characteristics, Chi-Square, Fisher's Exact, Independent Samples Mann-Whitney U and one Independent Samples t Test were used as follows: Chi-Square and Fisher's Exact tests were used, in accordance with the expected frequencies of the contingency tables, to cross-tabulate the frequencies of response in each category of preferred visual approach design against the categories of the following variables: region, helicopter transport capacity, helicopter cockpit generation, pilot's licence, pilot's gender, pilot's rank, and pilot's seniority. The Mann-Whitney U test was used against the following continuous variables: pilot's time as an offshore helicopter pilot, pilot's total flying hours in helicopters, pilot's IFR flying hours in helicopters, pilot's night flying hours in helicopters, pilot's number of night deck landings and pilot's suggested time gap for recency decay. Finally, the Independent Samples t Test was used to assess the parametric distributions of pilot's age across categories of preferred visual approach design.

The results of the statistical tests were all insignificant, which indicates that the choice of specific visual approach design follows the pilot's discretion only.

6. CONCLUSIONS

This paper has presented a structured survey methodology for identifying the critical flight phases of nighttime offshore helicopter flights. The survey overcomes the limitations of previous analyses with respect to low data quality, use of clustered phases of flight and low generalisability of results. This is of utmost importance given the current scenario of few accidents [6, 10], failed safety interventions based on accident data and reported incidents that cannot be trusted as precursors to accidents [8, 11].

Using a bespoke taxonomy which describes the human factors of the phases of nighttime helicopter flights, careful questionnaire design and piloting, collection of data from key subject matter experts and the study of the data's properties, multiple statistical hypotheses were tested. The results show that the phases flown by reference to external visual cues at high kinetic energy states (i.e., VH) cause the greatest concerns to pilots. Among such phases, the visual segment of instrument approaches is particularly problematic.

The analysis of risk levels experienced during this phase against the demographic variables failed to show any statistical significance. This is important as the sampling strategies of follow-up studies should be based on criteria other than the variables explored in this study. However, for the remaining VH phases (i.e., visual takeoff and go around), studies covering the operations of mid twin-turbine and glass cockpit helicopters should be given priority.

The perception of risk by mid twin-turbine (MT) helicopter pilots should be investigated further as such helicopters were previously associated with pilot-related and nighttime accidents [6, 10], higher risk ratings on the VH phases, but not with a perceived need for more frequent nighttime flying practice. Likewise, recency requirements for glass cockpits require further investigation as they appear to differ from those currently applied, which were established for pilots of legacy cockpits with analogic and mixed designs.

The study of a number of associated hypotheses revealed the 90-day pilot recency standard generally practiced in the industry is considerably longer than the average 28 day time gap suggested by the pilots for assured recency. A combination of the current standard of 3 cycles of takeoff/landing and the 28-day time gap suggested seems the best starting point for focused experimental research aimed to determine appropriate recency requirements. Experimental research should potentially involve pilot observation in high-fidelity flight simulators, as well as the analysis of real FDM data.

As safety interventions are put in place to address the per- phase-of-flight risks, it will be important to detect phases accurately and reassess the changes in associated risk levels. The statistical methods described in this paper will equally apply in the analysis of the risk levels attributed by pilots to more precisely-discriminated phases of flight.

7. METHODOLOGICAL CONSIDERATIONS

Surveys are necessarily based on subjective data such as experience, perceptions and information recollected from memory. Therefore, a degree of bias can never be discarded [11, 13]. Additionally, despite the attempted rigour of the sampling strategies, it could be argued that there was no true randomisation of the study sample. This stems from the fact that the pilots were volunteers [34]. While these are fair concerns, the simplicity of the questions and the focus on senior pilots with declarative knowledge and considerable professional experience are likely to have minimised any tendentious responses. These attributes should guide future survey attempts.

Representativeness can also be questioned. However, conversations with the collaborating individuals in each company revealed that the percentage representativeness was high in at least the North Sea countries, Brazil and Spain (i.e., over 30% of the population of 'senior pilots' in each region). Hence, the results can be expected to be a fair representation of the corresponding regions.

Finally, suboptimal sample sizes limited the analysis that could be undertaken, e.g., analysis by all regions of operations intended was not possible. However, given the ethical imperative to keep survey participation voluntary, low sample sizes often cannot be avoided. Therefore, the results should be interpreted with respect to such real-life constraints and dealt with as the first step towards evidence-based recommendations, rather than a fully validated study.

Acknowledgements

The Lloyd's Register Foundation and Santander Universities for funding, all the collaborating companies and pilots for the data, and the SMEs for their invaluable criticism of the methods.

References

1. Nascimento, F.A.C., et al., *A multistage multinational triangulation approach to hazard identification in night-time offshore helicopter operations.* Reliability Engineering & System Safety, 2012. **108**(0): p. 142-153.
2. Nascimento, F.A.C., A. Majumdar, and W.Y. Ochieng, *Investigating the truth of Heinrich's pyramid in offshore helicopter transportation*, in *Transportation Research Board 92nd Annual Meeting*. 2013, TRB: Washington, D.C.
3. Taber, M. and J. McCabe, *Helicopter Ditching: Time of Crash and Survivability.* SAFE Journal, 2006. **34**(1): p. 5-10.
4. Harris, F.D., E.F. Kasper, and L.E. Iseler, *U.S. Civil Rotorcraft Accidents, 1963 Through 1997*. 2000, NASA: Moffett Field, CA.
5. Jarvis, S. and D. Harris, *Investigation into Accident Initiation Events by Flight Phase, for Highly Inexperienced Glider Pilots.* International Journal of Applied Aviation Studies, 2008. **8**(2): p. 211-224.
6. Nascimento, F.A.C., A. Majumdar, and W.Y. Ochieng, *Helicopter Accident Analysis.* The Journal of Navigation, 2014. **67**(1): p. 145-161.
7. European Commission - Joint Research Centre. *ECCAIRS 4.2.8 Data Definition Standard: Event phases*. 2012 [cited 2013 10/04]; Available from: http://eccairsportal.jrc.ec.europa.eu/index.php/Products/83/0/.
8. Nascimento, F.A.C., A. Majumdar, and W.Y. Ochieng, *Incident reporting in offshore helicopter transportation*, in *38th European Rotorcraft Forum 2012*. 2012: Amsterdam, Netherlands.
9. U.S. Joint Helicopter Safety Analysis Team, *The Compendium Report: The U.S. JHSAT Baseline of Helicopter Accident Analysis, Volume I (CY2000, CY2001, CY2006)*. 2011, IHST.
10. Nascimento, F.A.C., A. Majumdar, and W.Y. Ochieng, *A 15-year multivariate analysis of worldwide offshore helicopter accidents*, in *AHS 69th Annual Forum and Technology Display*. 2013, AHS: Phoenix.
11. Nascimento, F.A.C., A. Majumdar, and W.Y. Ochieng, *Investigating the Truth of Heinrich's Pyramid in Offshore Helicopter Transportation.* Transportation Research Record: Journal of the Transportation Research Board, 2013. **2336**: p. 105-116.
12. Howson, D.A., *Research initiatives for improving the safety of offshore helicopter operations.* The Aeronautical Journal, 2006. **110**(1113): p. 14.
13. Nascimento, F.A.C., S. Jarvis, and A. Majumdar, *Factors Affecting Safety During Night Visual Approach Segments for Offshore Helicopters.* The Aeronautical Journal, 2012. **116**(1175): p. 303-322.
14. Nascimento, F.A.C., A. Majumdar, and S. Jarvis, *Nighttime approaches to offshore installations in Brazil: Safety shortcomings experienced by helicopter pilots.* Accident Analysis & Prevention, 2012. **47**(0): p. 64-74.
15. CAST/ICAO. *Phase of Flight - Definition and Usage Notes*. 2011 [cited 2012 10/06]; Available from: http://www.intlaviationstandards.org/Documents/PhaseofFlightDefinitions.pdf.
16. Teixeira, S., *Helicopter Flight Operational Quality Assurance (HFOQA): Development of HFOQA Analysis Software*. 2006, Embry-Riddle Aeronautical University: Daytona Beach, Florida.
17. Skybrary. *Flight Phase Taxonomy*. 2010 [cited 2012 05/06]; Available from: http://www.skybrary.aero/index.php/Flight_Phase_Taxonomy.

18. CAA, *CAP 382 - The Mandatory Occurrence Reporting Scheme*. 2005, CAA: Gatwick.
19. CAA, *CAA Paper 97009 - A Questionnaire Survey of Workload and Safety Hazards Associated with North Sea and Irish Sea Helicopter Operations*. 1997, CAA: London, UK.
20. Rollenhagen, C., et al., *The context and habits of accident investigation practices: A study of 108 Swedish investigators.* Safety Science, 2010. **48**(7): p. 859-867.
21. Smith, P.C. and L.M. Kendall, *Retranslation of expectations: an approach to the construction of unambiguous anchors for rating scales.* Journal of Applied Psychology, 1963. **47**(2): p. 149-155.
22. ICAO, *Safety Management Manual (SMM)*. 2012, ICAO: Montréal.
23. OGP, *Aircraft Management Guidelines*. 2008, International Association of Oil & Gas Producers: London, UK.
24. Nascimento, F.A.C., A. Majumdar, and W.Y. Ochieng, *Nighttime offshore helicopter operations – identification of contextual factors relevant to pilot performance*, in *4th International Conference on Applied Human Factors and Ergonomics (AHFE)*. 2012: San Francisco, USA.
25. Khatwa, R. and A.L.C. Roelen, *An Analysis of Controlled-Flight-into-Terrain Accidents of Commercial Operations, 1988 through 1994.* Flight Safety Digest, 1996. **17**(11-12): p. 166-212.
26. IHST. *The developing history of an accident-free future*. 2011 [cited 2013 04/11]; Available from: http://www.ihst.org/Default.aspx?tabid=1780&language=en-US.
27. Chipman, S.F., J.M. Schraagen, and V.L. Shalin, *Introduction to cognitive task analysis*, in *Cognitive task analysis*, J.M. Schraagen, S.F. Chipman, and V.L. Shalin, Editors. 2000, Lawrence Erlbaum Associates: Mahwah, NJ. p. 3-23.
28. Burns, R.P. and R.A. Burns, *Business Research Methods and Statistics Using SPSS*. 2008, London: SAGE Publications Ltd.
29. Frakes, M.A., K. High, and J. Stocking, *Transport Nurse Safety Practices, Perceptions, and Experiences: The Air and Surface Transport Nurses Association Survey.* Air Medical Journal, 2009. **28**(5): p. 250-255.
30. Field, A., *Discovering Statistics Using SPSS*. 3rd ed. 2009, London: SAGE.
31. van der Schaaf, T. and L. Kanse, *Biases in incident reporting databases: an empirical study in the chemical process industry.* Safety Science, 2004. **42**(1): p. 57-67.
32. van der Schaaf, T. and L. Kanse, *Checking for Biases in Incident Reporting*, in *Accident Precursor Analysis and Management: Reducing Technological Risk Through Diligence*, J.R. Phimister, V.M. Bier, and H.C. Kunreuther, Editors. 2004, National Academy of Engineering: Washington, D.C.
33. Harris, D., *Head-down flight deck display design*, in *Human Factors for Civil Flight Deck Design*, D. Harris, Editor. 2004, Ashgate: Aldershot.
34. Dery, M., et al., *Results and recommendations from the helicopter EMS pilot safety survey 2005.* Air Medical Journal, 2007. **26**(1): p. 38-44.

COPYRIGHT STATEMENT