Air Medical Journal 38 (2019) 198–201

Contents lists available at ScienceDirect

Air Medical Journal

journal homepage: http://www.airmedicaljournal.com/

Original Research

Evaluation of a New Helicopter Crew Transport Fatigue Assessment

Ryan Bushmaker, CFRN [1], Katherine Corey, MPA, MSN [1,*], Julie Dunn, MD [1], Trent Lalonde, PhD [2], Samantha Estrada, MS [2]

[1] *Medical Center of the Rockies, Loveland, CO*
[2] *University of Northern Colorado, Greeley, CO*

A B S T R A C T

Fatigue in air ambulance crews leads to decrements in performance and situational awareness that may contribute to aircraft accidents and patient care mistakes. Fatigue assessments completed by flight crews can give early warning when fatigue is accumulating. Countermeasures can then be implemented to improve performance and increase safety. No validated air ambulance fatigue assessment currently exists that incorporates transport-specific factors. The objective of this study was to validate a flight fatigue assessment that accounts for air transport–specific factors. Flight crewmembers from multiple air ambulance programs participated and completed assessments. Results were analyzed to determine if the assessment captured or predicted fatigue levels of crewmembers. When used to measure crewmember fatigue, the assessment was shown to consistently and reliably confirm accumulating fatigue and correlated with crew-reported levels of fatigue. A predicted fatigue scale was created to help crewmembers objectively identify their fatigue level. Used consistently, the transport fatigue assessment should increase awareness of accumulating fatigue. With awareness, crewmembers are better equipped to justify and take advantage of opportunities to mitigate their fatigue, increase crew coordination, enhance safety, and improve patient care.

In 2016, the National Transportation Safety Board added fatigue-related incidents to their most wanted list, and they continue to remain on their most wanted list for 2017 and 2018.[1,2] Commercial flight agencies have long recognized fatigue is a contributing factor in many air disasters.[3] Additionally, nurse and staff member fatigue increasingly is acknowledged as a significant threat to patient safety, prompting risk managers to adopt strategies to reduce fatigue caused by scheduling, overtime, and excessive workloads.[4] The Commission on Accreditation of Medical Transport Systems recommends assessing crewmembers' alertness and fatigue as part of a fatigue management system.[5]

A literature search looking for fatigue assessments specific to flight crews, shift workers, and medical professionals did not reveal any assessment tools that addressed fatigue factors uniquely affecting air ambulance crewmembers. Factors include altitude, patient acuity, transport time, and night vision goggles. Unfortunately, little work has been done regarding assessing fatigue in prehospital flight personnel, and no validated flight assessment tools for fatigue currently exist. In addition to factors specific to flight, other universal factors to account for are cumulative sleep debt, stress, nutrition status, and circadian rhythms. The goals of the study were 3-fold: 1) to develop a fatigue risk assessment scale; 2) to assess the reliability; and 3) to verify validity of the tool for predicting crewmember fatigue.

Methods

A literature search using PubMed and CINAHL was conducted looking for fatigue assessments specific to flight crews, and we could find no targeted assessment tools for this cohort. We then designed a transport fatigue assessment (TFA) that incorporated unique factors associated with air medical transport and included universal factors contributing to fatigue. After obtaining institutional review board approval, we conducted a series of assessments for phase 1 of the study. The assessment tool was then administered to crews at the following specific times during a crewmember's shift: at the beginning of the shift to assess baseline fatigue and after each transport to account for transport factors and provide awareness of accumulating fatigue. In phase 1, medical crewmembers from LifeLine and Sanford AirMed completed assessments and were also asked to complete an Epworth Sleepiness Scale (ESS) with each assessment.

Using the statistical tools described later, TFA was evaluated for internal consistency (the tendency of repeated questioning of similar

* Address for correspondence: Katherine Corey, MPA, MSN, Medical Center of the Rockies, 2500 Rocky Mountain Avenue, Loveland, CO 80538,
E-mail address: Katie.Corey@doane.edu (K. Corey).

1067-991X/$36.00

https://doi.org/10.1016/j.amj.2018.11.006

TRANSPORT FATIGUE ASSESSMENT					
Static Factors					
Done at the beginning of shift					
	0	**1**	**2**	**3**	**4**
How many hours did you sleep in the last 24 hours?	>8 hours	>7 hours	>6 hours	>5 hours	<5 hours
How well did you sleep in the last 24 hours?	Excellent	Good	Adequate	Fair	Poor
How many hours did you sleep in the last 72 hours?	> 24 hours	20-24 hours	18-19 hours	16-17 hours	< 16 hours
Rate your personal stress	None	Slight	Mild	Moderate	Severe
Are you ill?	No	Slight	Mild	Moderate	Severe
Dynamic Factors					
Complete Post Transport					
	0	**1**	**2**	**3**	**4**
What time of Day did your transport take place?	0701-1500	1901-2300	1731-1900	1501-1730 - 2301-0200	0201-0700
How many transports have you had this shift?	1	2	3	4	5 or more
How long have you been on duty?	<6 hours	6-11 hours	12-14 hours	15-17 hours	>18 hours
How many hours have you been awake?	<14 hours	14-15 hours	16-17 hours	18-19 hours	>20 hours
How long was your transport? (Depart-RTB)	<1 hour	1-2 hours	3-4 hours	5 hours	>6 hours
Did you use NVG's?	No				Yes
How many hours since you last ate?	4 hours	6 hours	8 hours	10 hours	>10 hours
How many hours since you last drank a non-caffeinated or alcoholic beverage?	2 hours	3 hours	4 hours	5 hours	>5 hours
	0	**1**	**2**	**3**	**4**

Figure 1. TFA Assessment Questions

participants to produce similar results) and for concurrent validity (the extent to which TFA measures what we intend to measure compared with an existing tool) by comparing TFA with the ESS. The ESS was chosen because of its popularity as a daytime sleepiness scale that has been accepted and validated through previous study.[6]

Phase 1 analyses of internal consistency included Cronbach alpha measures for both the TFA and ESS, for which values exceeding 0.70 on a scale of 0.00 to 1.00 are treated as evidence of strong internal consistency. Internal consistency was also assessed using an exploratory factor analysis. The tool was then refined by removing some questions proven irrelevant, confusing, or not helpful and by standardizing questions using a 5-point Likert Scale.

Concurrent validity was established by 2 linear mixed effects models, with one using the static TFA score as the outcome, the second using postflight TFA scores as the outcome, and both models including corresponding ESS scores as the predictor (and a random normal error to account for repeated observation of individuals

Table 1
Predicted Fatigue for Static Scores

Static score	4-7	8-12	13-15	16-21	-
Predicted fatigue	1	2	3	4	5

Table 2
Predicted Fatigue for Total Scores

Total score	11-17	18-23	24-32	33-48	49-58
Predicted fatigue	1	2	3	4	5

within a shift). In both models, the significance of the ESS as the predictor was taken as evidence of concurrent validity.

After refinement, the assessment was further studied in phase 2. Crews from LifeLine at Medical Center of the Rockies in Loveland Colorado USA, Air Link from Regional West Medical Center in Scottsbluff, Nebraska USA, and Health Net from Morgantown and Buckhannon in West Virginia USA, participated and completed the assessments. All crews worked either 12- or 24-hour shifts and were working at helicopter bases. Internal consistency of the TFA was reevaluated using Cronbach alpha and Guttman lambda, again using values of 0.70 to indicate strong internal consistency. Assessment questions (Fig. 1) were multiple choice, standardized into a 5-point Likert scale, and assigned different weights based on the answers that corresponded to increased fatigue. Crewmembers selected the assessment based on whether it was the beginning of their shift or after a transport. In addition to the assessment's baseline questions, 2 additional questions were added to the assessment. A shift length question was asked to see how fatigue might be affected by the length of the crewmember's shift. The second question asked the crewmember how fatigued they felt. Researchers felt it was important to try and correlate the crewmember's self-reported fatigue to what the assessment was predicting to see if a pattern could be established.

Validation for phase 2 was established by correlating beginning of shift (BOS) and postflight scores with self-reported fatigue measured on a scale from 1 (low fatigue) to 5 (high fatigue). Two ordinal multinomial logistic regression models were applied to establish this validity, both using self-reported fatigue (initially and after each flight) as the outcome of interest and one using BOS scores and the other using postflight TFA scores as the predictors. In each case, the probability of reporting the crewmember's level of fatigue was modeled using the TFA score, allowing the researchers to assess the strength of the association between TFA scores and self-reported fatigue and also to make fatigue predictions using TFA scores.

The ordinal multinomial models were used in a predictive way to identify the most likely level of fatigue associated with different ranges of static and total TFA scores. Internal consistency reliability analyses were performed in SPSS Version 20 (IBM Corp, Armonk, NY); validation models were fit using R Version 3.2.2 (Free Software Foundation, Inc., Boston MA).

Results

For phase 1, a total of 95 shifts were used for analysis, consisting of 95 BOS and baseline ESS scores followed by 63 postflight TFA and ESS scores. For this study, the internal consistency for the ESS was $\alpha = .72$ but was found to be 0.48 and 0.43 for the TFA BOS and postflight scores, respectively, which fall below the threshold of 0.70 for strong internal consistency. The exploratory factor analysis yielded 5 different factors associated with all of the individual items on the TFA. The items in these factors were examined along with the interitem correlation matrix, and, consequently, 7 unnecessary items were removed or reworded to increase consistency in user responses for phase 2. Items removed from the BOS assessment were time worked in the role and consecutive days worked while maximum altitude, acuity, scene call, low volume patient, and outside temperature were removed from the post flight questions of the TFA.

The phase 1 linear mixed model to establish validity of the BOS TFA scores showed a significant positive association between the ESS and the BOS TFA ($\hat{\beta} \approx .502,\ t^* \approx 4.421,\ P < .001$), indicating strong concurrent validity. The linear mixed model for the postflight scores showed a significant positive association between the ESS and postflight TFA scores ($\hat{\beta} \approx .167,\ t^* \approx 2.169,\ P \approx .034$), again suggesting evidence of concurrent validity.

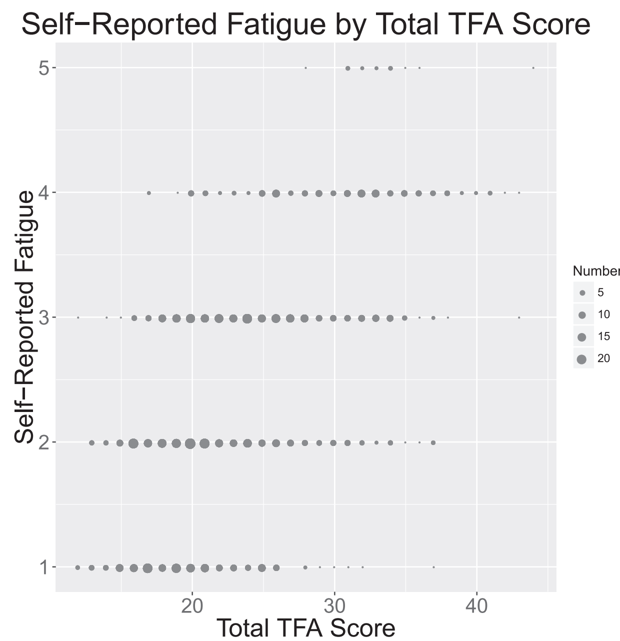

Figure 2. Self-Reported Fatigue Versus TFA Score

For phase 2, the data included 923 observations of the BOS TFA and 745 observations of postflight scores. A total of 10 BOS and 30 postflight assessment scores were excluded because of errors in entry of assigned unique participant identification. Cronbach alpha for the static assessment improved to 0.753, whereas Guttman lambda was 0.782; postflight assessment scores showed an increase to Cronbach alpha of 0.770 and Guttman lambda of 0.778. Overall, these results from phase 2 provide evidence of good reliability of the scores.

Figure 2 shows the observed relationship between the total TFA scores (static and postflight together) and self-reported fatigue. Higher TFA scores are associated with greater fatigue. Using the phase 2 logistic model to establish validity of the static assessment scores, there is evidence of a strong association between the BOS assessment score and self-reported fatigue at the beginning of the shift ($\hat{\beta} \approx .479,\ t^* \approx 18.050,\ P < .001$), indicating strong concurrent validity. The corresponding model for postflight scores also provided strong evidence of an association between postflight assessment scores and self-reported fatigue ($\hat{\beta} \approx .189,\ t^* \approx 15.190,\ P < .001$), again indicating concurrent validity. In both cases, the positive coefficient suggests that as the TFA score increases, the likelihood of reporting a higher level of fatigue increases.

Table 1 shows the predicted levels of fatigue according to the BOS TFA scores. Notice that the highest level of fatigue is not predicted and was not observed during static assessment. Table 2 shows predicted levels of fatigue according to the total TFA scores (static and postflight together).

Discussion

To increase safety, a specialized fatigue assessment is needed to assess air ambulance flight crews because of the unique environment in which they operate. Other sleepiness assessments currently exist including the popular ESS and the Stanford Sleepiness Scale, which were evaluated for their usefulness in the air ambulance environment. Although those sleepiness assessments are good for evaluating daytime sleepiness, they are not intended for use in shift workers, do not address aspects of working night shifts, and do not include fatiguing factors specific to flight crewmembers. When crews are flying, they are subjected to physiological stresses of flight, which potentiate fatigue. Those stresses include hypoxia, barometric pressure changes, thermal changes, decreased humidity, noise, and gravitational forces.[7] The longer crews are exposed to these stressors the more fatiguing their transport becomes.

In this study, prehospital flight teams from multiple sites were surveyed to evaluate a newly developed flight fatigue assessment. The TFA is the first validated flight crewmember fatigue assessment designed to overcome the shortcomings in other fatigue assessments by including transport fatiguing factors and considerations for circadian rhythms. The TFA is easy to complete and was usually completed faster once crews had been exposed and used the assessment a few times. Analysis showed the changes made after phase 1 were critical to user acceptance and response consistency. Accurate understanding of what the assessment questions are asking is important so crewmembers answer the questions consistently.

The TFA objectively assesses for fatigue. This is significant because fatigue self-reporting can be affected by physical activity, environment, and the human tendency to overestimate the current level of alertness and underestimate the current level of fatigue.[8] Our findings indicate there is a high likelihood that assessed fatigue levels with the TFA will correlate with the crewmembers' fatigue levels independent of their subjective feelings. When used, the assessment promotes insight and recognition of accumulating fatigue in individuals taking it and should help crewmembers identify when they are more likely to need to apply operational countermeasures to combat growing fatigue. An assortment of well-tested operational countermeasures is available to fatigued individuals and includes strategic napping, meditation, judicious caffeine use, exercise, rest breaks, time-outs, and uninterrupted sleeping.[9] This is assuming the assessment is completed in a spirit of truthful investigation into one's fatigue and not completed in a way to simply get a satisfactory score.

Helicopter emergency medical service programs may address fatigue differently depending on local policy and program understanding of fatigue and its implications. Although some programs may limit the number of consecutive days worked in a row, other may not. There are programs that allow crewmembers to sleep if not engaged in patient care or transport responsibilities, whereas other programs may require crews to assist in the hospital when they are not engaged in transport activities. Based on assessment reliability, program leadership and crewmembers should have confidence that TFA scores will accurately reflect fatigue levels and produce reliable results. If a validated and standardized fatigue assessment for flight crewmembers were accepted as an industry standard, with the majority of flight programs using the same fatigue assessment, benchmarking with other programs could become the norm, with the best fatigue mitigation practices being shared throughout the air medical transport industry.

We believe the TFA will provide an excellent framework for future research projects assessing fatigue and its effects on air medical providers. More research is needed that addresses shift duration in excess of 24 hours because none of the study groups scheduled staff longer than a 24-hour period. Future research is warranted to better understand the nuances of transport programs that routinely engage in ground, fixed wing, long-term, and international transports. There are industries outside air medical transport that may benefit from the TFA, which warrants future research and investigation. Specifically, those occupations that routinely use shift workers and who have high-risk jobs, such as firefighters, emergency medical technicians, nurses.

Conclusion

When fatigue levels are low, crewmembers are less susceptible to decrements in performance and therefore less prone to making patient care mistakes and better able to apply crew resource management concepts to increases team safety. A new TFA was developed and validated to help identify air ambulance flight crew fatigue, which accounts for unique stresses of flight and the human circadian rhythm. The TFA should prove to be a useful risk assessment tool and may contribute to a decrease in flight-related incidents, accidents, and patient care mishaps by increasing self-awareness of fatigue. Crewmembers and management, with increased awareness of accumulating fatigue, will be better equipped to combat it and intervene with mitigation strategies. This will aid crews in staying alert and being ready to respond to future transports safely.

Acknowledgments

The authors thank Josh Weiland for his participation, vision, and early support in the transport fatigue assessment study; Tracy Meyer and Jeffery White for their support and participation in the transport fatigue assessment study; and the crewmembers of LifeLine, Sanford, AirLink, and Health Net flight teams.

References

1. Most wanted transportation safety improvements. National Transportation Safety Board. https://www.ntsb.gov/safety/mwl/Pages/default.aspx. 2016. January 25, 2017.
2. 2017-2018 most wanted list. National Transportation Safety Board. https://www.ntsb.gov/safety/mwl/Pages/default.aspx. 2017. Accessed December 1, 2017.
3. Gregory K, Winn W, Johnson K, Rosekind M. Pilot fatigue survey: exploring fatigue factors in air medical operations. *Air Med J*. 2010;29(6):309–319.
4. Freeman G. Nurse fatigue a huge threat to patient safety. AHC Media. https://www.ahcmedia.com/articles/134803-nurse-fatigue-a-huge-threat-to-patient-safety-but-can-be-addressed. March 9, 2015. Accessed October 25, 2016.
5. Accreditation Standards of the Commission on Accreditation of Medical Transport Systems (CAMTS), 10th Edition. http://69.89.31.68/~camtsorg/wp-content/uploads/2017/06/10th_Standards_Free_-_021017.pdf. October 2015. Accessed January 2, 2016.
6. Smyth C. The Epworth Sleepiness Scale (ESS). Try This: Best Practices in Nursing Care to Older Adults. http://69.89.31.68/~camtsorg/wp-content/uploads/2017/06/10th_Standards_Free_-_021017.pdf. Accessed 1 October 2016.
7. Holleran R. ASTNA Patient Transport. 4th ed. St Louis, MO: Mosby; 2010.
8. Rosekind M, Gander P, Gregory K, et al. Managing fatigue in operational settings 1: physiological considerations and countermeasures. *Behav Med*. 1997;21:157–165.
9. Caldwell J, Mallis M, Caldwell J, Paul M, Miller J, Neri D, Fatigue countermeasures in aviation. Aviat Space Environ Med 2009;80:29–59