# Fatigue Incident Antecedents, Consequences, and Aviation Operational Risk Management Resources

Megan B. Morris; Megan D. Wiedbusch; Glenn Gunzelmann

**BACKGROUND:** Flight crew fatigue is an important factor in aviation, leading organizations to implement fatigue risk management programs to reduce risk. The U.S. Air Force Air Mobility Command (AMC) has implemented the Aviation Operational Risk Management (AvORM) program to aid mission schedulers and flight crews in mitigating flight risks and identifying appropriate levels of risk. The AvORM program uses a scheduling tool and underpinning biomathematical fatigue model. This study examined self-reported fatigue-related incidents within AMC, which provides some indirect and anecdotal evidence as to the effectiveness of the scheduling tool.

**METHODS:** Archival data from the AMC Aviation Safety Action Program (ASAP) Safety Reporting System was examined. Report content themes were created through an inductive approach in terms of fatigue prevalence, antecedents, and consequences.

**RESULTS:** Fatigue was estimated as a factor in 4% of the reports. The two most commonly referenced fatigue antecedents were associated with mission/duty length and mission scheduling/planning factors. Factors associated with aircraft operation violations were the most cited consequences of fatigue. Fatigue was almost twice as likely to be reported as a secondary rather than primary contributing factor. Aircrew reported both positive and negative aspects of AvORM resources in mission planning and fatigue mitigation.

**DISCUSSION:** Examination of ASAP reports suggests that fatigue is a contributing factor to safety incidents. Although the AvORM program highlights potential flight risks by utilizing a scheduling tool built upon an underlying biomathematical fatigue model, human fatigue continues to impact safety, suggesting an ongoing need for improved fatigue risk management and mitigation.

**KEYWORDS:** fatigue, aviation, scheduling tool, biomathematical fatigue model.

**Morris MB, Wiedbusch MD, Gunzelmann G. *Fatigue incident antecedents, consequences, and aviation operational risk management resources*. Aerosp Med Hum Perform. 2018; 89(8):708–716.**

Flight crew fatigue is a significant concern in aviation, in both civilian and military contexts.[6,7] Researchers estimate that fatigue contributes to 4 to 8% of aircraft incidents, potentially costing millions of dollars in damage for a single incident and leading to fatal outcomes.[7] In addition, these incidents can cause delays and cancellations of subsequent flights and missions, disrupting the effectiveness of the organization. As a result, fatigue risk management is a critical consideration for organizations in the aviation industry to ensure the safety of individuals, reduce unnecessary and undesirable costs, and maintain organizational effectiveness.

The U.S. Air Force Air Mobility Command (AMC) has implemented the Aviation Operational Risk Management (AvORM) program to aid mission schedulers and flight crews in identifying, mitigating, and managing flight risks such as fatigue. One component of the AvORM program uses a scheduling tool and underpinning biomathematical fatigue model to estimate effectiveness for specific mission profiles. The current study examines safety related incidents regarding fatigue within AMC. Specifically, we examine reports from the AMC Aviation Safety Action Program (ASAP), a voluntary safety reporting system for aircrew members that serves as a repository of data related to safety issues that AMC uses to inform policy and interventions to maximize safety in operations.

---

From the Air Force Research Laboratory, Wright-Patterson AFB, OH, Ball Aerospace & Technologies Corp., Fairborn, OH, and Oak Ridge Institute for Science and Education, Belcamp, MD.

This manuscript was received for review in October 2017. It was accepted for publication in May 2018.

Address correspondence to: Megan Morris, Ball Aerospace & Technologies Corp., Bldg. 852, 2620 Q St., Wright-Patterson AFB, OH 45433; megan.morris.1.ctr@us.af.mil.

DOI: https://doi.org/10.3357/AMHP.5019.2018

We use reports submitted to the ASAP system to identify incidents that are related to fatigue within the database, as well as antecedents and consequences that were reported along with the fatigue incidents. Analysis of the safety incident reports will allow us to examine the severity of the threat fatigue poses to operational safety and effectiveness, and will provide some indirect and anecdotal evidence as to the effectiveness of the AvORM scheduling tool.

Despite the importance of fatigue in many occupational fields (e.g., military, transportation, and medical), there is not a single, agreed-upon definition of fatigue due to its multidimensional nature.[3,14,31] Research suggests that fatigue has at least as many as five dimensions, consisting of general fatigue, mental fatigue, physical fatigue, sleepiness, and lack of motivation or activity.[3] Aviation-related organizations have various fatigue definitions (for examples see the Federal Aviation Adminstration's[12] and the International Civil Aviation Organization's[19] fatigue management systems). Antecedents of fatigue vary depending on the definition, but often include factors such as lack of sleep, circadian phase, and physical or mental workload. Regardless, they all commonly refer to fatigue resulting in reduced alertness and mental and physical performance. These diminished abilities cause a safety concern in aviation operations. In addition, fatigue can be acute or can accumulate over time (i.e., chronic fatigue),[9] increasing the complexity of its effects. In spite of the varying conceptualizations of fatigue among fields and within aviation, it is widely acknowledged as an important factor that has contributed to numerous incidents and accidents. For that reason, fatigue remains a focal point in flight crewmember safety.

Many studies have examined the role of fatigue in accidents, as well as fatigue in general, across civilian, commercial, and military aviation.[7,25] Fatigue has considerably contributed to incidents within the U.S. Air Force. The Air Force Safety Center estimated that fatigue played a role in 7.8% of Air Force Class A mishaps.[23] In addition, fatigue was found to contribute to 25% of Air Force night tactical fighter Class A accidents between 1974 and 1992.[30] Recently, while examining spatial disorientation in Class A U.S. Air Force mishaps, researchers reported that 6.66% of mishaps were related to fatigue.[29] Flight crewmembers are especially vulnerable to fatigue given the operational environments experienced by these individuals. Research has shown factors such as low air flow and light levels, vibration, restricted movement, background noise, sustained vigilance, and automation, are common in flight crew environments and are contributors to fatigue.[4]

Studies suggest that fatigue often acts as one of several contributors to aviation incidents.[7,8] In most cases, fatigue is not identified as the primary cause of aircraft accidents and mishaps.[23] While this is true, it is important to understand that fatigue often serves as a compounding factor. If other environmental factors are present, such as increased workload, sustained vigilance, or deviations from normal procedures, fatigue can interact with these factors and increase the likelihood of an incident.[33]

Within aviation, several factors have been shown to influence fatigue. Studies on the U.S. Air Force have identified circadian rhythm disruption, sleep disruption, operational tempo, and scheduling issues as primary contributing factors.[27] Miller and Melfi[27] suggested poor scheduling as the main cause of fatigue in the Air Force. However, poor scheduling contributes to other factors that impact fatigue, such as circadian disruption from changes in multiple time zones or poor sleep quality from scheduled early morning flight duty.

Fatigue can result in a variety of symptoms for flight crewmembers. Research has suggested that common symptoms include negative impacts on short-term memory, reaction time, motivation, work efficiency, and increases in anxiety, irritability, sleepiness, risk-taking behavior, and errors of omission.[5,7,27] As a result, fatigue has been associated with decreased performance,[28] which can then lead to myriad incidents, such as aircraft operation deviations.[21]

Given the potentially catastrophic effects of fatigue within the aviation sector, it is important that effective fatigue risk mitigation processes are in place. Researchers have identified several possible strategies to battle flight crew fatigue, including educating flight crew about fatigue antecedents and consequences, incorporating strategic naps and breaks into flight schedules, the use of stimulants, and standard reporting criteria for incidents, among others.[2,7]

One strategy focuses on scheduling as a contributor to flight crew fatigue. Researchers have argued that when developing the flight, mission schedule planners/schedulers need to keep circadian and homeostatic factors in mind, including how these factors can interact to create fatigue.[7] The Air Force has taken this into consideration and has sponsored the development of the Fatigue Avoidance Scheduling Tool® (Fatigue Science, Vancouver, BC, Canada),[16] based on the Sleep, Activity, Fatigue, and Task Effectiveness™ (Institutes for Behavior Resources, Baltimore, MD)[18] model. The scheduling tool is used in the mission planning process to generate assessments of fatigue risk based on work/rest schedules. The model takes into account circadian rhythm, homeostatic regulation, sleep/wake schedules, and mission location information to predict performance effectiveness.

Once appropriate information is provided, alternative schedules can be compared based on fatigue predictions in the form of predicted performance effectiveness through graphs and tables.[11,18] On the effectiveness graphs, performance effectiveness is denoted as a line indicating level of performance throughout the schedule, roughly corresponding to percent optimal. The graph is sectioned into three effectiveness zones with green, yellow, and red representing various thresholds of performance effectiveness depending on the task at hand. The graph also includes a default dotted line at 77.5% performance effectiveness (within the yellow zone) that suggests the implementation of fatigue countermeasures when performance effectiveness drops below this value. This line can be adjusted depending on the task (e.g., piloting an aircraft vs. filling out documentation). The effectiveness tables include performance effectiveness scores, as well as other performance scores (e.g., lapse index, reaction time) and fatigue factors (e.g., chronic sleep debt, hours awake), for specific time points within the schedule. The tool can also be used to examine possible antecedents of fatigue retrospectively.[11,18]

Research has provided evidence for the validity of the biomathematical fatigue model and scheduling tool coupling as a predictor of fatigue. Research performed at the 2002 Fatigue and Performance Modeling Workshop in Seattle demonstrated the biomathematical model used in AvORM to have the least error in predicting subjective ratings of fatigue as well as objective vigilance performance compared to other fatigue models.[32] Other research comparing observed cognitive performance during fatigued states to model predictions provides additional support of its effectiveness.[16,18] Additional research has shown model predictions to correlate with accident risk related to human factors,[17] providing further support for its validity.

AMC's AvORM program uses the scheduling tool to produce AvORM Mission Effectiveness Graphs. The AvORM program helps highlight risk and provides the opportunity to mitigate or accept risk during mission planning/scheduling/execution. The program also detects risks that occurred after execution of the mission and provides mitigation plans based on those risks. Given the utilization of the scheduling tool and biomathematical fatigue model within the AvORM program, we were interested in examining the current status of fatigue in AMC aviation operations. Specifically, we examined the presence of fatigue in self-reported safety concerns as well as the antecedents and consequences associated with these occurrences.

## METHODS

### Procedure

The study protocol was determined as Not Human Subject Research by the Air Force Research Laboratory Institutional Review Board under the common rule (32 CFR 219), 17-123 AFRL IR NHSR Determination. We examined the ASAP Safety Reporting System database scoreboard for reports involving fatigue. ASAP is a web-based reporting tool that allows flight crewmembers to voluntarily self-report hazards, errors, and risks. As a result, it is not required that incidents be reported to this system. For example, Air Force Instruction 11-202V3[1] states "incidents involving damage to aircraft, personal injury, or intentional disregard of orders or instructions, whether reported to ASAP or not, shall be reported to a Flight Safety Officer (FSO) as soon as possible (T-0)." So, not all entries in the database reflect incidents that result in negative outcomes. The reporting is nonpunitive and nonidentifying, allowing for users to openly report concerns that otherwise might not be reported. System users can either submit an ASAP Report or a Fatigue Report, or they can view a scoreboard that consolidates all of the reports in a published, de-identified format.

To submit a report the submitter must designate their user group (e.g., Flight Crew, Maintenance and Logistics, All Other Specialties) and then designate an appropriate subgroup (e.g., for the Flight Crew user group, the Major Weapon System category must be identified). Report formats are tailored for each group. Commonly, the user will list the associated aircraft and provide a description of the event. In addition, the submitter can include a title on the report, departure and landing airports, reactions to the incident, suggestions concerning the incident, and can upload any supplemental material that might be helpful to the report (e.g., AvORM effectiveness graphs). Additional details can be provided about the incident, but not all of these details are published on the ASAP consolidated scoreboard. Once the report is submitted, the ASAP analyst team de-identifies the report and publishes the report to the scoreboard, where it is available to all users. The analyst team personnel also provide status and resolution entries updating the submitter on the status of the report, including information that has been gathered and actions that have been taken.

At the time of analysis, the reporting system database consisted of 2541 reports over a period of 7 yr and 4 mo. When viewing the scoreboard the user can view specific category entries (e.g., "Fatigue") separately by checking the appropriate boxes and searching for those entries. Examining the "Fatigue" related entries, we added 46 reports to our initial dataset. Additionally, the search component of the database allows users to search for reports with specific keywords. We used the following key words: fatigue, fatigued, tired, tiredness, tiring, sleep, sleepy, sleepiness, drowsy, drowsiness, exhausted, and exhausting, and added the resulting report instances to the initial dataset. We then went through the dataset reports to examine whether the report content was appropriate for our analysis. We used the following criteria to include a report in the final dataset:

1) The report must include human fatigue (or a description that fits human fatigue) as an antecedent of an incident, or the incident itself, in the report.
2) The report must delineate a specific occurrence of fatigue, not a summary of multiple occurrences or a discussion of possible future incidents regarding fatigue.
3) The fatigue mentioned in the report must be experienced by the submitter or one of their crewmembers, as opposed to speculation about someone not associated with their crew.

This resulted in the identification of 103 reports for our dataset.

### Statistical Analysis

To conduct our qualitative analysis of the fatigue-related content of the reports, we used each report, including the description, reaction, suggestions, and resolution of the report, as our unit of measurement. Given that the reports allow for open-ended responses and the option to include reactions and suggestions along with descriptions, it is possible that multiple antecedents, consequences, and characteristics of fatigue were reported by the submitter. As a result, the commonly used interrater reliability metric, Cohen's kappa, which involves mutually exclusive codes, was not appropriate for this analysis. We used an extension of Cohen's kappa formula, the proportional overlap procedure by Mezzich et al.,[26] to conduct the interrater reliability analysis. This extension of Cohen's kappa allows for multiple codes per case by allowing for partial agreement. A common metric for kappa assessment is: slight agreement, 0.00–0.20; fair agreement, 0.21–0.40; moderate agreement, 0.41–0.60; substantial agreement, 0.61–0.80; and almost perfect agreement, 0.81–1.00.[22]

It should be noted that kappa, along with Mezzich's extension of kappa, does not take into account the total number of possible codes that can be used for a given report. As the number of possible codes and the complexity of the entries increases, one would expect the kappa values to be generally decreased. The nature of the reports in the study raise this possibility. At the same time, the reports created an opportunity to use a hierarchy of emerging themes. Eccleston et al.[10] employed this strategy and used Mezzich's extension to examine interrater reliability analysis for interviews. We used a similar technique and generated themes with an inductive approach. Each resulting theme had a title, definition, and code number. Initially the two raters read through all the reports to identify possible themes. The possible themes were categorized into more general themes, assigned a title, and given definitions for coding.

To increase the validity of the coding, the raters decided to rate all of the codes independently in an iterative fashion. This process increases confidence that both raters' perceptions were taken into account for each report, rather than individual subjectivity and biases being present in the coding. The raters used the following process: selections of reports were independently coded using the coding system, Mezzich's kappa was calculated to examine interrater agreement, disagreements among the raters were discussed, the raters came to a consensus on the final coding, and appropriate updates were made to the coding sheet. The first three iterations involved a random selection of 20 reports and the final iteration involved the remaining reports. An example of a single coding for a report might be "133", which is "1 Antecedents, 13 Sleep Quality, 133 Base" (i.e., poor sleep quality from base accommodations as an antecedent of fatigue).

The first, second, third, and last iteration resulted in a Mezzich's kappa of 0.18, $t(19) = 4.33$, $P < 0.001$, $\mathrm{Se}(\kappa) = 0.04$; 0.34, $t(18) = 6.74$, $P < 0.001$, $\mathrm{Se}(\kappa) = 0.05$; 0.46, $t(19) = 12.09$, $P < 0.001$, $\mathrm{Se}(\kappa) = 0.04$; and 0.43, $t(41) = 14.03$, $P < 0.001$, $\mathrm{Se}(\kappa) = 0.03$, respectively. This suggested a general increase in rater agreement through each iteration and resulted in a moderate level of interrater reliability.[22] Many of the disagreements in coding involved choosing different subthemes. For example, there was confusion over the use of "billeting" by the system users. Some users used this term to refer to off-base accommodations (e.g., hotel), while others used the term to refer to other military base accommodations (e.g., a naval base). This caused the raters to use different subthemes for the "poor sleep quality" theme ("billeting" vs. "base"). Note that the second and last iteration involved one less report. The raters came to the consensus that one of the entries in both iterations was most likely a duplicate of another entry made by the same submitter or a different submitter from the same crew.

## RESULTS

After completion of the qualitative coding, we estimated that fatigue was a contributing factor in about 4% of the report entries in the database. We then computed frequency counts for the final codes derived from the fatigue-related reports (see **Table I**, **Table II**, and **Table III**). The two most commonly referenced antecedents of fatigue were associated with the Mission/Duty Length and Mission Scheduling/Planning factors. Mission/Duty Length was referenced in 32.87% of fatigue-related antecedents. The following are some example entries from aircrew that highlight this issue:

> *"Long duty day resulted in the aforementioned mental and emotional fatigue symptoms."*

> *"Poor off-base crew rest facility was a corntibuting [sic] factor, but don't lose sight off [sic] the main problem: 24-hour flight duty periods."*

> *"Fatigue - We were a basic crew lading [sic] with one hour remaining in our Flight Duty Period which included 10 hrs of flight time."*

> *"Several factors contributed to this, including crew fatigue from two 18+ hour duty days with minimum ground time/crew rest…"*

> *"…a long duty day contributed directly to this near mishap."*

Mission Scheduling/Planning was referenced in 25.18% of fatigue-related antecedents. For example, aircrew members stated:

> *"This mission was a perfect storm of LEGAL planning causing EXTREME cumulative fatigue."*

> *"Safety of flight was almost called due to compounded crew fatigue on leg 5/8 of a planned 5-day mission."*

> *"If the Mission Effectiveness graph goes below 60% something should be done early in the planning process to mitigate risk from fatigue."*

> *"The crew had also been getting up at 3:00 am every morning and working full 12 our [sic] tactical duty days. I believe this led to the crew being very fatigued."*

> *"While the crew implemented an appropriate work-rest cycle on the aircraft the next day, it could not combat the lack of sleep caused by wakefulness at [redacted] and crew performance suffered throughout the entire 24 hour flying period the following day."*

> *"...fatigue undoubtedly played a role at the end of a fairly intense training sortie well outside most crewmember's normal circadian rhythms."*

In terms of consequences of fatigue, factors associated with Incorrect Aircraft/Vehicle Operation were the most cited, representing 47.90% of consequences. Altitude deviation was the most commonly reported violation. For example:

**Table I.** Codes, Descriptions, Frequencies, and Proportions of Antecedents.

| CODE | DESCRIPTION | *F* | %SS | %S | %C | %T |
|---|---|---|---|---|---|---|
| 1 Antecedents | Statement of factors that contributed to fatigue | 143 | | | | 39.83 |
| 11 Changes | Changes in the mission or duty day from delays or extensions | 27 | | | 18.88 | 7.52 |
| 111 Delays | Mechanical failures, weather, and other environmental factors causing delays | 21 | | 77.78 | 14.69 | 5.85 |
| 112 Extensions | New stops added, called back for additional work, and other related extensions | 6 | | 22.22 | 4.20 | 1.67 |
| 12 Length | Length of mission or duty, or near end of mission or duty | 47 | | | 32.87 | 13.09 |
| 121 Long | Long planned mission or duty day, 12+ hours | 34 | | 72.34 | 23.78 | 9.47 |
| 122 Near End | Near the end of the mission | 13 | | 27.66 | 9.09 | 3.62 |
| 13 Sleep Quality | Poor crew rest due to environmental factors such as noise, light, air conditioning issues, etc. | 21 | | | 14.69 | 5.85 |
| 131 Aircraft | Sleeping in aircraft | 2 | | 9.52 | 1.40 | 0.56 |
| 132 Billeting | Sleeping in nonmilitary accommodations like a hotel | 4 | | 19.05 | 2.80 | 1.11 |
| 133 Base | Sleeping in base accommodations | 11 | | 52.38 | 0.37 | 3.06 |
| 14 Scheduling | Planning/scheduling in terms of sleep cycles and circadian rhythm | 36 | | | 25.18 | 10.03 |
| 141 Nonstandard Duration | Nonstandard crew rest duration or odd times | 12 | | 33.33 | 8.39 | 3.34 |
| 1411 Too Short | Crew rest was too short | 0 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1412 Too Long | Crew rest was too long | 9 | 75.00 | 25.00 | 6.29 | 2.51 |
| 1413 Odd Rest | Crew rest was at an odd time | 3 | 25.00 | 8.33 | 2.10 | 0.84 |
| 142 Odd Flight Time | Odd flight times effecting circadian rhythm | 23 | | 63.89 | 16.08 | 6.41 |
| 1421 Time Zone | Time zone change | 5 | 21.74 | 13.89 | 3.50 | 1.39 |
| 1422 Night/Day | Duty at night/very early morning | 15 | 65.22 | 41.67 | 10.49 | 4.18 |
| 15 Task Demands | Task demands created overload or limited completion time | 3 | | | 2.10 | 0.84 |
| 151 Task Saturation | Overload due to too much work or multiple tasks | 3 | | 100.00 | 2.10 | 0.84 |
| 152 Compressed Time | Fast paced work or limited time to complete task | 0 | | 0.00 | 0.00 | 0.00 |
| 16 Physical Limits | Physical overload or limitations reached | 9 | | | 6.29 | 2.51 |
| 161 Physical Activity | Increased physical activity | 1 | | 11.11 | 0.70 | 0.28 |
| 162 Dehydration | Physical dehydration from insufficient water consumption | 1 | | 11.11 | 0.70 | 0.28 |
| 163 Extreme Heat | Extreme heat on airplane or outside | 4 | | 44.44 | 2.80 | 1.11 |
| 164 Altitude | High altitude resulting in hypoxia (fatigue is a side-effect) | 2 | | 22.22 | 1.40 | 0.56 |

*F* = frequency; %SS = percentage of sub-sub-category; %S = percentage of sub-category; %C = percentage of category; %T = percentage of total.

> *"Due to fatigue (11 hours into the duty day), channelized attention on the approach briefing, and two different radios simultaneously broadcasting, the aircraft descended through FL220 to FL215. At FL215, the PF noticed the deviation and immediately corrected the aircraft altitude to FL 220. The aircraft deviated from FL220 for no more than 10 seconds."*

Fatigue was almost twice as likely to be reported as one of many factors contributing to incidences reported, rather than the primary contributor. For example, in several reports aircrew mentioned several contributing factors:

> *"Fatigue, workload and CRM were also insidious contributing factors."*

> *"This incident also highlights the potential consequences of a series of compounding factors: 1) combat offloads … 7) fatigue undoubtedly played a role at the end of a fairly intense training sortie well outside most crewmember's normal circadian rhythms."*

Whereas others specifically flagged fatigue as the primary factor:

> *"The crew is declaring safety of flight due to fatigue among all crewmembers."*

A small number of fatigue-related reports (4.74%) referenced objective fatigue evaluations from the AvORM effectiveness graph and the AvORM worksheet Fatigue Risk score. These quotations highlight mission scheduling/planning issues, suggesting a need for increased emphasis on effectiveness predictions and risk scores in the planning process. For example, one aircrew member suggested:

> *"Schedule missions referencing the forecast crew effectiveness."*

The reports suggest that the AvORM program can have a positive impact on operational decision making to improve safety, including fatigue risk management. For example, one submitter wrote:

> *"The crew alerted and I discussed each person's level of fatigue. After each crewmember individually scored their own Fatigue levels based upon the AMC ORM guide, I determined that we were in the "High" ORM category for fatigue… I contacted the TACC/DOO [Tactical Air Control Center/Director of Operations] and explained our situation. After some coordination with the TACC/Senior it was determined that the crew would re-enter Crew Rest and set for an LFA [Legal for Alert] 24-h from the original alert time."*

The submitter stated that it "…was a successful case of ORM-scoring being applied to TACC decision making."

Two other examples:

> *"The Aircraft commander consulted with all crew positions and it was determined that the crew could no longer safely operate the aircraft due to fatigue. Aircraft commander reworked the ORM worksheet attached and came up with a score of 27, in the*

**Table II.** Codes, Descriptions, Frequencies, and Proportions of Consequences.

| CODES | DESCRIPTION | *F* | %SS | %S | %C | %T |
|---|---|---|---|---|---|---|
| 2 Consequences | Statement of outcomes that fatigue contributed to | 119 | | | | 33.15 |
| 21 Mission | Mission/sortie/training cancelled or delayed, thinking about or called safety of flight | 25 | | | 21.01 | 6.97 |
| 211 Mission Cancelled | Mission was cancelled | 1 | | 4.00 | 0.84 | 0.28 |
| 212 Safety of Flight Called | Safety of flight was called | 7 | | 28.00 | 5.88 | 1.95 |
| 213 Safety of Flight Almost Called | Thinking about calling safety of flight, or that it was a possibility | 10 | | 40.00 | 8.40 | 2.79 |
| 214 Delay for Crew Rest | Mission delayed for crew rest | 6 | | 24.00 | 5.04 | 1.67 |
| 215 Sortie/Training Cancelled | Sortie/training was cancelled | 1 | | 4.00 | 0.84 | 0.28 |
| 22 Communication | Communication errors among crewmembers, tower, and any other relevant individuals | 26 | | | 21.01 | 2.82 |
| 221 Misunderstanding | Misunderstood commands or thought individuals said something different | 10 | | 38.46 | 8.40 | 2.79 |
| 2211 Crew | Between aircrew members | 1 | 10.00 | 3.85 | 0.84 | 0.28 |
| 2212 Tower | Between aircrew members and tower | 9 | 90.00 | 34.62 | 7.56 | 2.51 |
| 222 Missed | Completely missed communication | 9 | | 34.62 | 7.56 | 0.02 |
| 2221 Crew | Between aircrew members | 0 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2222 Tower | Between aircrew members and tower | 8 | 88.89 | 30.77 | 6.72 | 2.23 |
| 223 Other Communication | Other communication issues that do not fit two previous categories | 6 | | 23.08 | 5.04 | 1.67 |
| 2231 Crew | Between aircrew members | 1 | 16.67 | 3.85 | 0.84 | 0.28 |
| 2232 Tower | Between aircrew members and tower | 3 | 50.00 | 11.54 | 2.52 | 0.84 |
| 23 Operations | Incorrect aircraft/vehicle operations | 57 | | | 47.90 | 15.88 |
| 231 Flap Speed Deviation | Flap speed deviations | 6 | | 10.52 | 5.04 | 1.67 |
| 232 Altitude Deviation | Altitude deviations | 11 | | 19.30 | 9.24 | 3.06 |
| 233 Taxi Incursion | Incursions that occurred exclusively while taxiing | 4 | | 7.02 | 3.36 | 1.11 |
| 234 Landing without Clearance | Crew landed without tower clearance | 7 | | 12.28 | 5.88 | 1.95 |
| 235 Course Deviation | Course deviation during flight | 7 | | 12.28 | 5.88 | 1.95 |
| 236 Other | Other operational incidents | 22 | | 38.60 | 18.49 | 6.13 |
| 24 Stimulants | Use of stimulants such as caffeine or go pills | 1 | | | 0.84 | 0.28 |
| 25 Burnout | Physical or mental burnout | 1 | | | 0.84 | 0.28 |
| 26 Non-Event | Explicitly stated that no adverse events occurred or everything went smoothly | 9 | | | 7.56 | 2.51 |

*F* = frequency; %SS = percentage of sub-sub-category; %S = percentage of sub-category; %C = percentage of category; %T = percentage of total.

> *high category. He felt the safest course of action was to return to crew rest.”*

> *“The crew came up with a [sic] ORM score of 14 with 6 points for crew fatigue. This moderate total score with a high single risk factor… After further discussion with the crew and the squadron DO, it was determined that this was a safety of flight call.”*

In contrast, some reports showed that the AvORM processes have limitations:

> *“The ORM process in place does not accurately reflect how crewmembers feel or how rested they are. Despite having HIGH ORM scores in the categories of personal ORM and crew rest/fatigue, the overall ORM score was in the MEDIUM range.”*

These reports suggest that the scheduling tool and the underlying biomathematical fatigue model as well as the process applying these tools can be improved. Situations might arise when operations deviate from the original plan, possibly resulting in different sleep schedules, or time zone changes, among other fatigue-related factors that have changed from the original derived schedule. In these cases, fatigue scales are frequently not updated in AvORM, which can hide impacts on crew fatigue. For example, one airmen wrote:

**Table III.** Codes, Descriptions, Frequencies, and Proportions of Fatigue Characteristics.

| CODE | DESCRIPTION | *F* | %SS | %S | %C | %T |
|---|---|---|---|---|---|---|
| 3 Fatigue Contribution | Statement that fatigue was the primary or one of many contributing factors to incident | 80 | | | | 22.28 |
| 31 Primary | Primary contributing factor to incident | 29 | | | 36.25 | 8.08 |
| 32 One of Many | One of many contributing factors to incident | 51 | | | 63.75 | 14.21 |
| 4 Objective Fatigue Level (AvORM) | References to FAST™ graph within AvORM or AvORM Fatigue Risk Score | 17 | | | | 4.74 |
| 41 Yellow Effectiveness | Yellow graph area | 0 | | | 0.00 | 0.00 |
| 42 Orange/Red Effectiveness | Orange/Red graph area | 8 | | | 47.06 | 2.23 |
| 43 Fatigue Risk Score | High AvORM worksheet Fatigue Risk score | 9 | | | 52.94 | 2.51 |

*F* = frequency; %SS = percentage of sub-sub-category; %S = percentage of sub-category; %C = percentage of category; %T = percentage of total; AvORM: Aviation Operational Risk Management; FAST: Fatigue Avoidance Scheduling Tool.

> *"I didn't have access to Aviation ORM but I knew the ORM had to be high since before he had changed it we were flying into the Yellow zone and now since it was Augmented it had to be bad."*

In another instance, ASAP system personnel responded to one entry as follows:

> *"While we have no doubt the crew was experiencing fatigue, it was difficult to correlate with the ME [Mission Effectiveness] graph and AvORM fatigue database due to the extensive mission changes and no updates to AvORM to capture some of these changes."*

## DISCUSSION

Examination of the ASAP Safety Reporting System database suggests that fatigue is a contributing factor to many safety incidents that were reported. We note that as we went through the database we noticed duplicates of some fatigue-related reports under different Air Force community categories. Additionally, other instances of fatigue might have been mentioned in reports that we were not able to find due to other keywords being used to describe fatigue, or fatigue could possibly be inferred from the context of the report. It is difficult to produce an exact percentage of reports that contained instances of human fatigue. Based on the number of entries we found and duplicates, we estimate that this figure is around 4% (out of 2541 reports). This is similar to other estimates[7] that have been reported. The proportion of accidents involving fatigue might be higher since the ASAP Safety Reporting System is voluntary and crewmembers are asked to report accidents involving personal injury, death, or aircraft damage to a wing safety officer. As a result, some situations involving fatigue are likely missing from the ASAP Safety Reporting System.

The analysis suggested that Mission/Duty Day Length and Mission Scheduling/Planning were the most prominent antecedents of fatigue. Past research[15,27] has also suggested that these are primary causes of fatigue. Other notable antecedents that were identified paralleled past research, such as poor sleep quality.[27] Although Mission/Duty Day Changes was also a notable antecedent, it should be noted that these changes sometimes resulted in an increased mission or duty day length.

One of the most prevalent consequences of fatigue, either as a primary or contributing factor, were incorrect aircraft operations. This is consistent with other research examining fatigue-related incidents.[21] Incidents such as altitude, flap, clearance, and taxi deviations, which usually occur during descent, approach, and landing, were prevalent. This finding is consistent with past research[20,24] examining fatigue related incidents. This is a significant concern given the potentially catastrophic consequences of incorrect operations in this environment.

Although AvORM effectiveness graphs and fatigue risk scores were referenced only in a small number of the entries, the content of the entries suggest that these are very important resources for both mission planning and aircrew. Some of the references to the AvORM effectiveness graphs and fatigue risk scores were positive, suggesting that the graphs and scores were useful in assessing fatigue risk and were used to make important mission decisions such as declaring safety of flight and extending crew rest. This suggests that aircrew perceive AvORM resources are beneficial and that they have value in mission planning and fatigue mitigation during missions.

However, some entries suggested that effectiveness graphs and fatigue risk scores were not used to their full potential in the mission planning process, and that in some cases the graphs and scores were inaccurate during mission execution. Mission changes often lead to inaccurate or outdated mission effectiveness and fatigue risk scores. Aircrew are often not able to examine updated effectiveness graphs because access to the graphs requires an authorized computer with an Internet connection, which is not always available to aircrew. In addition, the current biomathematical fatigue model does not take into account certain factors such as stimulant use. Stimulants such as caffeine were not frequently mentioned in the reports; however, the actual usage could be greater given that caffeine is found in many beverages that are a part of normal, daily consumption (e.g., coffee and soda), which might not be regarded by aircrew in fatigue mitigation processes. Another limitation with the current biomathematical model implementation in AvORM is that it does not generate graphs that are individually tailored to each aircrew member, but rather reflect an overall performance effectiveness level for the entire aircrew. These limitations should further direct efforts to increase the validity of the biomathematical model within AvORM and ensure that these predictions are being effectively utilized in the mission planning process and in aircrew fatigue mitigation during the mission.

The current study has limitations that should be kept in mind. Readers should be cautious when interpreting the results of these incident reports as they were all self-report from various crew positions across the Air Force. Depending on the crewmember, they might have differing views as to the contribution of fatigue in an incident (for example see Lyman and Orlady[24]). As a result, these reports mainly involve subjectively identified factors; however, some references to factors such as AvORM effectiveness graphs and fatigue risk scores were objectively identified. In addition, there is also subjectivity in the coding of the two raters for this project. Other raters might have coded these reports differently, especially in regard to the subthemes.

In the current study, we examined fatigue incidents primarily associated with aircrew members. It is important to note that there are other personnel involved with aircraft in addition to aircrew members who can suffer from fatigue and, as a result, contribute to incidents, such as air traffic controllers[13] and line maintenance crews.[34] Similar to aircrews, these personnel might also benefit from a fatigue risk management program using a scheduling tool and underlying biomathematical fatigue model to combat fatigue.

Despite efforts to curtail and manage crew flight fatigue, the AMC ASAP Safety Report System suggests that fatigue is still a contributor to safety related incidents. Our analysis of the fatigue-related reports suggests that long missions/duty days and mission planning are the primary contributors to fatigue. Incorrect aircraft and vehicle operation issues such as altitude deviations were the most common consequence of fatigue. Results suggest that although the scheduling tool within the AvORM program is a beneficial tool that can potentially be used to effectively manage fatigue, fatigue risk management must continue to evolve to consider the broader context in which people are operating. Given the 24/7 nature of military operations, it is impossible to completely eliminate fatigue and fatigue risk from operations. Minimizing risk through careful scheduling and appropriate restrictions on duty hours is a critical component of effective fatigue risk management. In addition, however, attending carefully to the causes and consequences of incidents is critical to identifying how fatigue risk manifests in operations. Self-report systems like ASAP provide a significant opportunity to gather data that does not involve damage to systems, injury, or loss of life.

## ACKNOWLEDGMENTS

The views presented in this paper are those of the authors, and do not represent the official position of the United States Air Force or Department of Defense. We would like to thank the U.S. Air Force Air Mobility Command for providing guidance in the access of the Aviation Safety Action Program database and for their assistance in manuscript editing. This research was funded in part with support from the Defense Health Agency, Joint Program Committee on Military Operation Medicine, Fatigue Mechanisms Working Group (88ABW; Cleared 10/17/2017; 88ABW-2017-5025).

*Authors and affiliations:* Megan B. Morris, Ph.D., Ball Aerospace & Technologies Corp., Fairborn, OH; Megan D. Wiedbusch, B.S., Oak Ridge Institute for Science and Education, Belcamp, MD; and Glenn Gunzelmann, Ph.D., Air Force Research Laboratory, Wright-Patterson AFB, OH.

## REFERENCES

1. Air Force. Flying operations. General flight rules. Air Force; 2016. Air Force Instruction (AFI) 11-202, Vol. 3. [Accessed 6 Dec. 2016]. Available from: static.e-publishing.af.mil/production/1/acc/publication/afi11-202v3_accsup_i/afi11-202v3_accsup_i.pdf.
2. Åkerstedt T. Consensus statement: fatigue and accidents in transport operations. J Sleep Res. 2000; 9(4):395.
3. Åkerstedt T, Knutsson A, Westerholm P, Theorell T, Alfredsson L, Kecklund G. Mental fatigue, work and sleep. J Psychosom Res. 2004; 57(5):427–433.
4. Battelle Memorial Institute & JIL Information Systems. An overview of the scientific literature concerning fatigue, sleep, and the circadian cycle. Washington (DC): FAA Office of the Chief Scientific and Technical Advisor for Human Factors; 1998. [Accessed 18 Nov. 2016]. Available from: https://www.vcockpit.de/fileadmin/dokumente/presse/2003/BatelleReport1998.pdf.
5. Bennett SA. Flight crew stress and fatigue in low-cost commercial air operations – an appraisal. Int J Risk Assess Manag. 2003; 4(2/3):207–231
6. Caldwell JA. Crew schedules, sleep deprivation, and aviation performance. Curr Dir Psychol Sci. 2012; 21(2):85–89.
7. Caldwell JA. Fatigue in aviation. Travel Med Infect Dis. 2005; 3(2):85–96.
8. Command Flight Surgeon. Fatigue in naval aviation. Pensacola (FL): Naval Aviation Schools Command; 2005. [Accessed 21 Nov. 2016]. Available from https://www.public.navy.mil/NAVSAFECEN/Documents/aviation/SAS/fatigue_in_naval_aviation.pdf.
9. Craig A, Cooper RE. Symptoms of acute and chronic fatigue. In: Smith AP, Jones DA, editors. Handbook of human performance. San Diego (CA): Academic Press; 1992:289–340.
10. Eccleston P, Werneke U, Armon K, Stephenson T, MacFaul R. Accounting for overlap? An application of Mezzich's K statistic to test interrater reliability of interview data on parental accident and emergency attendance. J Adv Nurs. 2001; 33(6):784–790.
11. Eddy DR, Hursh SR. Fatigue avoidance scheduling tool (FAST™), phase II SBIR final report, part 1. Dayton (OH): NTI, Incorporated; 2006. Report No: AFRL-HE-BR-TR-2006-0015. [Accessed 1 May 2017]. Available from: http://www.dtic.mil/dtic/tr/fulltext/u2/a452991.pdf.
12. Federal Aviation Administration. Fatigue risk management systems for aviation safety (AC 120-103A). Washington (DC): Federal Aviation Administration; 2013. [Accessed 6 Dec. 2016]. Available from: https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_120-103A.pdf.
13. Gander P. Fatigue management in air traffic control: the New Zealand approach. Transp Res F Traffic Psychol Behav. 2001; 4(1):49–62.
14. Gartner WB, Murphy MR. Pilot workload and fatigue: a critical survey of concepts and assessment techniques. Washington (DC): National Aeronautics and Space Administration; 1976. Report No: NASA/Technical Note D-8365. [Accessed 22 Nov. 2016]. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19770004731.pdf.
15. Goode JH. Are pilots at risk of accidents due to fatigue? J Safety Res. 2003; 34(3):309–313.
16. Hursh SR, Balkin TJ, Miller JC, Eddy DR. The fatigue avoidance scheduling tool: Modeling to minimize the effects of fatigue on cognitive performance. SAE Transactions Journal of Aerospace. 2004; 113(1): 111–119.
17. Hursh SR, Raslear TG, Kaye AS, Fanzone JF Jr. Validation and calibration of a fatigue assessment tool for railroad work schedules, summary report. Washington (DC): Federal Railroad Administration; 2006. Report No: DOT/FRA/ORD-06/21 [Accessed 1 May 2017]. Available from: https://www.ble-t.org/pr/pdf/dot_fra_ord_0621.pdf.
18. Hursh SR, Redmond DP, Johnson ML, Thorne DR, Belenky G, et al. Fatigue models for applied research in warfighting. Aviat Space Environ Med. 2004; 75(3):A44–A53.
19. International Civil Aviation Organization. Manual for the oversight of fatigue management approaches, 2nd ed. International Civil Aviation Organization; 2016. Doc 9966. [Accessed 28 Apr. 2017]. Available from: https://www.icao.int/safety/fatiguemanagement/FRMS%20Tools/Doc%209966.FRMS.2016%20Edition.en.pdf.
20. Kowalsky NB, Masters RL, Stone RB, Babcock GL, Rypka EW. An analysis of pilot error-related aircraft accidents. Albuquerque (NM): Department of Aerospace and Environmental Medicine Lovelace Foundation; 1974. Report No: NASA Contractor Report 2444. [Accessed 28 Nov. 2016]. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740018321.pdf.
21. Krueger GP, Jones YF. U.S. Army aviation fatigue-related accidents. Fort Rucker (AL): U.S. Army Aeromedical Research Laboratory; 1978. Report No: USAARL Report No. 79-1. [Accessed 28 Nov. 2016]. Available from: https://www.usaarl.army.mil/TechReports/79-1.pdf.
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33(1):159–174.
23. Luna T. Fatigue in context: USAF mishap experience [Abstract]. Aviat Space Environ Med. 2003; 74(4):388.
24. Lyman EG, Orlady HW. Fatigue and associated performance decrements in air transport operations. Moffett (CA): NASA Ames Research Center; 1981. Report No: NASA Contractor Report 166167. [Accessed 28 Nov. 2016]. Available from: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19810014163.pdf.
25. Mallis MM, Banks S, Dinges DF. Aircrew fatigue, sleep need and circadian rhythmicity. In: Salas E, Maurino D, editors. Human factors in aviation, 2nd ed. Burlington (MA): Academic Press; 2010:401–436.
26. Mezzich JE, Kraemer HC, Worthington DRL, Coffman GA. Assessment of agreement among several raters formulating multiple diagnoses. J Psychiatr Res. 1981; 16(1):29–39.

27. Miller JC, Melfi ML. Causes and effects of fatigue in experienced military aircrew. Brooks City-Base (TX): Air Force Research Laboratory; 2006. Report No: AFRL-HE-BR-TR-2006-0071. [Accessed 28 Nov. 2016]. Available from: http://www.dtic.mil/dtic/tr/fulltext/u2/a462989.pdf.
28. Neville KJ, Bisson RU, French J, Boll PA, Storm WF. Subjective fatigue of C-141 aircrews during Operation Desert Storm. Hum Factors. 1994; 36(2):339–349.
29. Poisson III RJ, Miller ME. Spatial disorientation mishap trends in the U.S. Air Force 1993-2013. Aviat Space Environ Med. 2014; 85(9):919–924.
30. Ramsey CS, McGlohn SE. Zolpidem as a fatigue countermeasure. Aviat Space Environ Med. 1997; 68(10):926–931.
31. Tiesinga LJ, Dassen TWN, Halfens RJG. Fatigue: a summary of the definitions, dimensions, and indicators. Nurs Diagn. 1996; 7(2):51–62.
32. Van Dongen HPA. Comparison of mathematical model predictions to experimental data of fatigue and performance. Aviat Space Environ Med. 2004; 75(3):A15–A36.
33. Van Dongen HPA, Hursh SR. Fatigue, performance, errors, and accidents. In: Kryger MH, Roth T, Dement WC, editors. Principles and practice of sleep medicine, 5th ed. St. Louis (MO): Elsevier Saunders; 2010:753–759.
34. Wang TC, Chuang LH. Psychological and physiological fatigue variation and fatigue factors in aircraft line maintenance crews. Int J Ind Ergon. 2014; 44(1):107–113.