Fatigue Risk Management System

Best Practice Control of Fatigue Risk Throughout the Crew Management Process

This document summarizes and explains best practices for limiting and reducing fatigue risk when constructing and operationalizing airline crew rosters. The guidance provided assumes the use of crew planning optimizers and a bio-mathematical model (BMM) to support the evaluation of fatigue risk in crew pairings and rosters.

Please note: Guidance provided assumes that the reader has reviewed the "Best Practices for Quantification of Fatigue Risk" [1] which separately describes fatigue risk and metrics for quantification such as AFR and NFR.

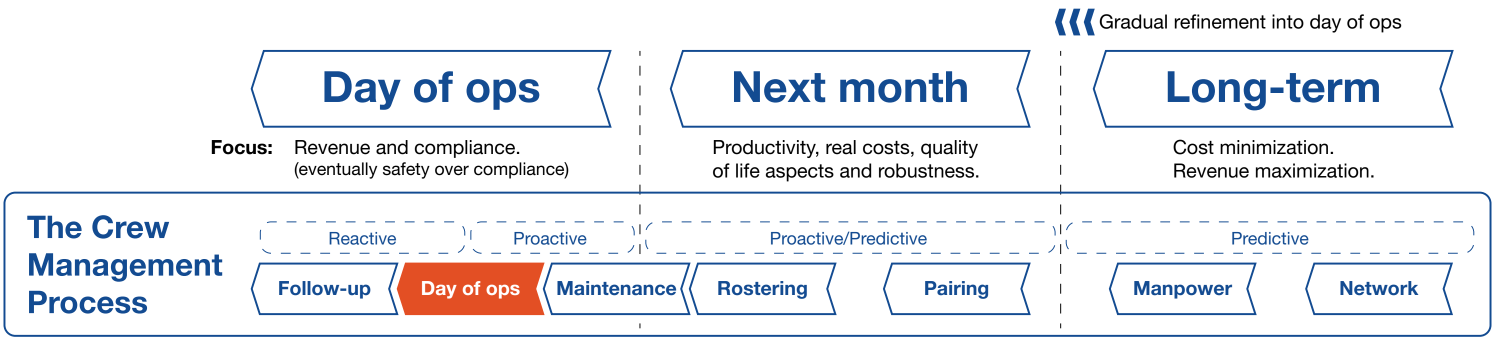

Figure 1. The crew management process.

The Crew Management Process

The crew management process is fairly complex for most airlines and covers many separate steps that are revisited repetitively as the flight schedule or crew availability changes over time. The main steps are depicted above with future time to the right, with less detailed long-term planning results, which are gradually refined moving to the left, and eventually 'meet reality' in day of operation. In this process flow, there are predictive, proactive and reactive fatigue risk management activities as presented in the figure. But let's start from the beginning, with network planning.

Network Planning and the Flight Schedule

Fatigue risk develops primarily from poor time of day, extended wakefulness and a prior sleep debt, as dictated by our physiology. Given this, the baseline fatigue risk for an operator is, to a great extent, defined by the flight schedule. By avoiding flights where two-pilot operation is performed, requiring high-performance (for example, an approach and landing at a difficult airport, 03:30-05:30 AM body clock time and the crew has been operating throughout the night) the risk profile becomes less challenging. In an operation with multiple crew bases in different time zones, or with crew operating sequences exceeding a couple of days, such that acclimatisation occurs, it becomes much harder to estimate crew body clock time when scrutinizing the flight schedule. In these cases, there is a need to first build crew pairings (see further below) before assessing risk with any accuracy.

Some flights should be questioned when setting the flight schedule:

- Will this operation match our "risk appetite"?
- Should we pass up this revenue opportunity?
- Even if we can fly and fulfill regulation?

Too often the fatigue risk impact is underestimated or ignored when constructing the flight schedule. When striving for maximizing utilization and revenue, the reasoning may be along the lines of, "If we can operate within regulation, the crew planning department can manage fatigue risk later somehow." It may be that the flight is re-timed to come in just below the FTL limits, making it possible to operate on two pilots instead of three. But the regulatory FTLs are over-simplifications of something much more complex.

There is no good way of rostering a two-pilot flight with a report time at home base of 17:00 and arrives 11 hours later into a difficult airport with a night landing. It is unlikely that crew will find time to increase their alertness, in the prior afternoon, and will be impacted by long wakefulness and poor time of day. The flight duty, regardless of how it is planned, will be a fatiguing one. The seemingly efficient two-pilot plan, when setting the flight schedule, ends up either being one with high risk or eventually augmented. Analysis can be done in this stage on new or changed routes using a BMM.

Best practice is to produce crew pairings, and even rosters, on upcoming flight schedule periods (so called budget runs) well ahead of time to identify fatigue hot spots using a BMM, potentially resulting in a modified schedule. Risk is then quantified and compared to the normal levels operated in order to identify upcoming negative trends and issues. Such budget runs are also used for verifying the base establishment, ensuring no upcoming crew deficits, per rank and qualification, down to a daily level. A crew deficit may otherwise lead to elevated risk as it will require denser rosters with less margin, in turn leading to unstable rosters.

Fatigue reports, on certain flights, may be relevant to consider if they are sufficiently frequent and if those flights, repeatedly end up in the same context. In short-haul or cargo operations, it is difficult to gather fatigue reports per flight number as the context varies constantly; flown as the fourth sector on the fourth day, and another time as the second sector on the first day.

Network Planning Summary

- Question flights
- Re-time critical phases
- Budget runs
- Quantify and track risk over time
- Model new routes
- Qualify and use operational feedback

So what about pairing construction, surely that must be much easier?

Pairing Construction

A crew pairing is a sequence of flights for a set of crew starting and ending at a crew home base. It is an efficient building block for crew not yet known by name. These pairings are constructed to ensure all flights in the flight schedule receive the necessary crew complement and to respect crew establishment numbers per base. It also considers other hard constraints, such as rules from the regulator and the union. The overall goal is to find a feasible solution that minimizes an overall cost (or objective function) that reflects real costs, productivity, solution robustness and quality-of-life aspects for crew.

Best practice operators allow for a BMM to feed into the cost function so that fatiguing flights are seen as more expensive to the optimizer, effectively reducing overall fatigue risk while the solution is being built. The BMM is thus used as a counter-force to balance the resulting force from management and crew as seen in Figure 2, keeping the solution away from unsafe conditions [8].

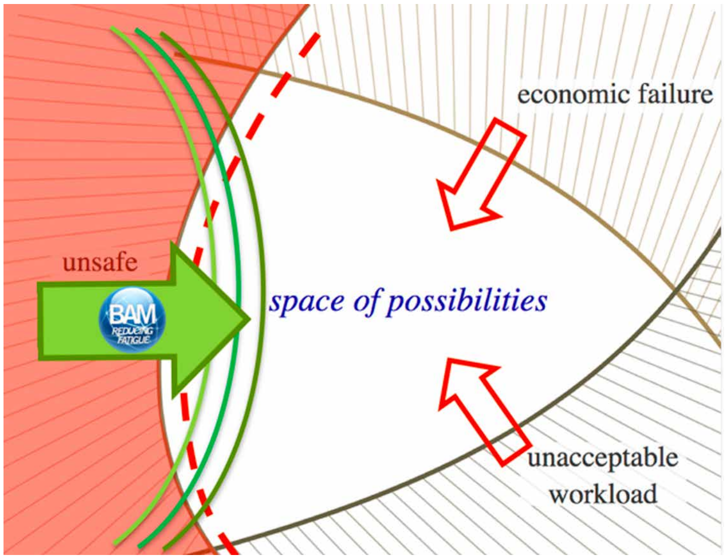

Figure 2. Using a BMM as a protective counter-force in Rasmussens' model Space of Possibilities

Additionally, some operators include a hard rule for transparency, ensuring a lower bound on the predicted alertness, such as a minimum alertness at top of descent of 1250. Alternatively, a rule is written for the lowest, or the average, level during the flight. Such rules, although providing simplicity, are ineffective in reducing or limiting the more important overall fatigue risk. Perhaps articulated best at a recent health and safety conference, "Trying to reduce overall fatigue risk by only using a rule on minimum alertness level, is like trying to reduce your weight by only deciding to skip cheese cake." A more comprehensive approach is needed to make a real impact.

Some operators, that lack a BMM capable of influencing the solution during optimization, instead measure fatigue risk after the fact when the solution is produced and then perform manual rework. This approach prolongs lead times, requires additional effort, and also moves the solution away from the optimum by adding duty days, layovers and/or positioning flights. Ironically, when the planning process is prolonged, less accurate data is used as the flight schedule evolves, further adding to elevated costs.

Trying to reduce overall fatigue risk by only using a rule on minimum alertness level, is like trying to reduce your weight by only deciding to skip cheese cake.

It is important to notice that the pairings in a solution may look better than they really are. This as they may, when sequenced on rosters, generate a significantly different (higher) fatigue risk from what can be seen in the pairings alone. This rosterability aspect can be controlled by either informing the BMM about an estimated roster context (if the BMM allows for that), or by letting the pairing optimizer construct longer working periods; pairings containing fictive layovers at home base. These working periods are then cut up at home base, before being rostered.

The hard constraints, which the optimizer must respect, may also drive fatigue risk. Best practice is to vary parameter settings, producing what-ifs in order to learn which rules may be counter-productive and subject for reformulation. There is methodology available for doing so in a systematic way [2]. Example: a shorter layover or rest may reduce fatigue risk on subsequent flights. Some of these rules can be reformulated by the airline, while some would need approval from the authority.

Fatigue risk may quietly build up over time, which is why best practice includes tracking fatigue risk per fleet/rank/base and over time to identify trends in the data. NFR is one good metric for this but AFR should also be used to pay attention to the overall risk. Efforts should be focused on structural problems that are both frequent and severe in order to obtain the best risk reduction per invested dollar.

It is further possible, but often quite difficult, to limit fatigue risk by introducing additional rules. The human physiology is complex, and rules will always be simplifications. There is a risk that a good optimizer will find creative ways of dealing with the new rules when searching for efficiency, potentially causing even higher risk in other cases that are not caught by the rules. New rules will also limit the optimizer in terms of improving the overall objective, and often the rules are better at limiting crew efficiency than reducing fatigue risk [9]. Best practice here includes stress-testing the rule set by applying an inverted penalty on fatigue risk in a what-if scenario to identify loopholes. More about this in the referenced material below [6].

Rather than only limiting and suppressing fatigue risk across the operation when building pairings, it is also possible to modify the cost experienced by the optimizer with a multiplier elevating it for night landings, difficult airports, etc. This allows for positioning the alertness of crew where it matters the most, based on the experience of the organization.

Strategic buffers, or margins to regulatory or other hard rule limits, are very important for the stability of the rosters later in day of operation. Without well placed buffers in the pairings, delays and other disruptions will snowball, causing a multitude of roster changes, some of those at short notice to crew. Tools such as Jeppesen Calibration use advanced analytics on disruption history guiding the decisions on optimal buffering. Overly protective buffers may be very expensive.

Pairing Summary

- Suppress overall risk with a BMM
- Improve buffers (up/down)
- Position alertness
- Introduce rules (carefully, and stress-test rule set)
- Consider building working periods
- Vary settings and learn through what-ifs
- Quantify and track AFR/NFR per fleet/rank/base over time
- Start pairing publication work late, on accurate data, and plan fast
- Qualify and use operational feedback

So what about rostering these pairings?

Roster Construction

Rostering comes in a few different flavors, mainly weighted fair share and strict seniority rostering. The strict seniority regime, common in North America, puts a lot of the sequencing of the pairings into rosters on the crew. The crew, for the most part, approach this in seniority order via a bidding system. The company has limited influence on how pairings are combined; if the combination is legal it should go on the roster. In the weighted fair share regime, normal outside of North America, the company influence is more direct and there is no constraint dictating that a sequence will need to be given to crew, even if asked for.

Staying with the weighted fair share regime, best practice is to suppress risk using a BMM during rostering optimization by influencing the objective function, just as for pairing construction. Everything else valued by the company will also be present as cost elements in the objective function. There will, for example, be an award in the optimizer for granting a bid, so a crew member asking for a fatiguing combination of pairings will have two elements opposing each other in the objective function - an award if the bid is granted, against a penalty for a fatiguing combination. Best practice, recognizing that a model has no detailed knowledge about crew individuals (the model is under-informed), and that the model predicts averages for a population (limited accuracy on one individual), is to tune the bids to have a higher influence on the outcome. Should it be the case that crew continuously asks for (verified) fatiguing combinations - that is dealt with as an exception and addressed through information/training given to the individual.

Crew are trusted with the aircraft and the wellbeing and safety of the passengers. They should also be trusted, until proven differently, that they know more about their personal physiology, life situation, and fatigue sensitivities/resilience than a model developed for the average of a population, given that they are properly informed and trained.

It is crucial for crew to have influence over their rosters in order to keep the real (not the predicted) alertness levels in check. Due to rather significant individual differences between crew in how/when they experience sleepiness, they may withhold on information which is not considered or accounted for by the model. They know if they are a morning or evening person, how sleepy they typically become when staying up overnight, their commute time, family situation, habits, etc. They also do more than work and sleep; they have (of course!) a life outside work that drives a need for wakefulness at certain times such as family commitments, hobbies, etc. Due to this, it is crucial, and a best practice, to let crew express their desires through bids and requests on their upcoming roster before being produced.

Crew may place bids and assign them different priorities: end before 19:00 on Thursdays at home base, short pairings, layover in GOT max 2 per month, fly with crew 12345. There could also be bids such as: avoid fatiguing flights, or avoid night work that may influence the optimization. A request on the other hand, is more of a guarantee that will be fulfilled if granted such as: day off on Dec 25. The requests come in limited supply, tightly controlled by the airline to ensure feasibility when solving the rostering problem (staffing all flights). Bids and requests are typically collected over a web interface prior to planning. The requests are then preassigned and locked to the roster, whereas the bids are up to the optimizer to fulfill as best as possible influenced by an award element in the cost function.

Another type of influence is the type of crew contract. Best practice is to allow crew some options for switching base or fleet, or changing to work part-time. This allows crew to further match their workload to their personal life situation. A misalignment between the two may lead to higher fatigue risk as sleep typically suffers when the work/life balancing act becomes too hard.

Contrary to the pairing step, it is possible to individualize predictions from a BMM in rostering. Since the crew member is now known by name, it is possible to collect and inform the model on diurnal type, habitual sleep length, commute times, etc. Doing so may be quite political, but will deliver a more accurate prediction in the cases where crew deviate from the generic assumption otherwise applied.

Another good practice is to use what is called fairness functionality to distribute risk over the crew (Figure 3). The overall solution will contain a certain amount of fatigue risk, quantified by, for example, the AFR metric. By dividing the total AFR with the total available full-time-equivalent (FTE) number of crew, a full time target is set. Each crew will then get their fair share determined by a personal target being their availability to take on production, multiplied with the full-time target. This is typically done automatically before the optimizer starts. Any deviation from the personal target will feed into the optimizer objective function as a penalty for poor fairness - effectively driving the optimizer to come up with an even and fair distribution over crew. This is often done also by taking history into account; crew that came in low last planning period will get some of the deficit back as an elevated target in this planning period. There are also penalties allowing for separating fatiguing duties in time as an extra precaution. Instead of rostering two challenging night duties five days apart, the penalty makes the optimizer prefer placing them a couple of weeks apart. Most BMMs do not take long-term effects into account as there is shortage of data quantifying these effects, so a separate mechanism, like the one described is useful.

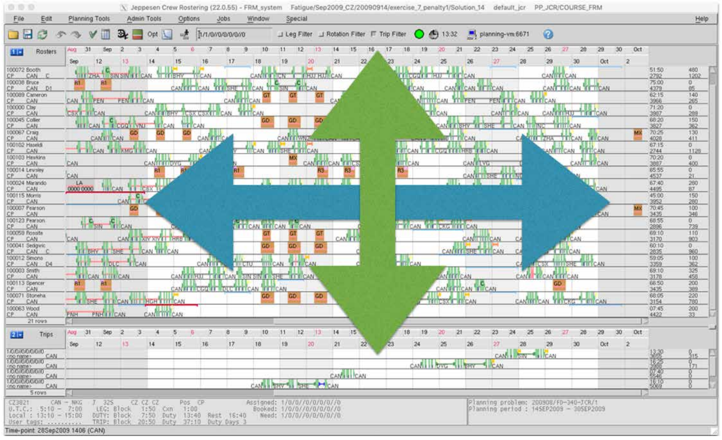

Figure 3. Distributing risk over crew vertically (green), and within crew horizontally (blue).

The comment in the pairing section on buffers is valid also for rosters. The rules that govern how the pairings may be combined should contain buffers reducing the risk for snowballing changes in day of operation. An example could be that a duty must end at least 90 minutes before midnight prior to a single day off. Otherwise, a delay would connect duty days to extend beyond the limit for maximum consecutive days, causing a pairing to be re-assigned to another crew with short notice. Buffers should be questioned regularly as it is much easier to add them than spot costly over-buffering. Again, methodology and tools are available to tune a buffer strategy in a data-driven way.

An additional robustness aspect is the creation, methodology and levels of planned reserve and standby duties. A good process in this regard delivers standbys into day of operation that are well positioned, usable, and in a ratio correlating to the production and sickness levels. It can be a complex process, but absolutely crucial for delivering robustness into the steps that follow.

Regarding the introduction of new rules to a rule set, the same comment as for the pairing step is valid here - verify that it is a real problem, exercise caution in the formulations so that crew efficiency is protected by using what-ifs, quantify the effect on risk and stress-test the rules to find loopholes. Use available methodology [2].

And again, best practice includes quantifying and tracking risk development (AFR/NFR) per fleet/base/rank over time. It is recommended to track the development weekly, as the weekly numbers match the periodicity of the flight schedule.

The rostering step ends with a roster publication to crew where expectations are set on how they will operate for the next period, typically a calendar month.

Rostering Summary

- Suppress overall risk with a BMM
- Improve buffers (up/down)
- Position alertness
- Introduce rules (carefully!)
- Consider building working periods
- Vary settings and learn through what-ifs
- Quantify and track AFR/NFR per fleet/rank/base over time
- Start publication work late on accurate data and plan fast
- Consider collecting more personal information to predict with higher precision

What happens next?

Roster Maintenance

The roster maintenance phase in the crew management process typically covers the time from roster publication to roughly 48 to 72 hours prior to day of operation. The work is mostly manual and focused on maintaining the good properties of the published rosters as information changes. The work in this phase includes attending to illegalities and crew shortages as crew go on long-term illness, lose qualifications or for other reasons, where availability changes. Changes to the flight schedule also have a big impact; new inserted charter flights, fleet swaps, etc.

A flight changing fleet may have a smaller impact on cabin crew in case they are cross-qualified, but are likely to require re-planning of the pilots. Crew may also initiate changes referred to as trip trading. Two or more crew may want to swap the production between them for a number of reasons - change of plans, not happy with their published roster, etc. Best practice includes enabling and supporting this process, in recognition of crew influence being a good thing as earlier explained - provided the alterations are also done with fatigue risk impact in mind.

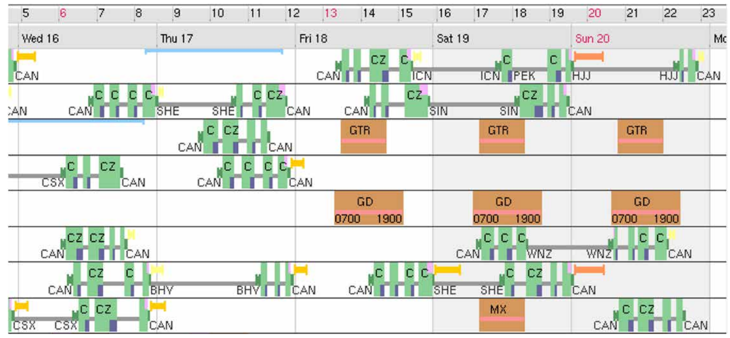

Figure 4. Fatigue markers updating live in a GUI (yellow and amber markers) reflecting predicted alertness.

With all of these changes being made to the rosters manually, it is good to have supporting functionality providing a visual risk indicator in the user interface creating awareness; for example fatigue markers (Figure 4) being displayed when risk thresholds are being exceeded. It is also recommended to track the development of AFR and NFR using a BMM.

It is good practice to avoid company-initiated changes to the rosters as much as possible, in particular closer to day of operation as it may leave crew less time to prepare. With good functionality to locate assignable crew and other decision support, these can be kept to a minimum. Many airlines have a possibility to "buy" free days from crew when there is a deficit, something that should be avoided or focused to streaks of days off that are longer than two, in order to avoid single days off which would leave less margin for recovery.

One big task for roster maintenance is to deliver the right robustness into day of operation. This is done by closely monitoring the daily usable standby level (per type of standby) and making sure what is handed over (taking place daily) can cope sufficiently with disruption.

Roster Maintenance Summary

- Take advice from BMM in rosters updates
- Allow for trip trading
- Safeguard that robustness is maintained for day of ops
- Don't change more than needed, especially at the end of maintained period
- Quantify and track AFR/NFR change per fleet/rank/base over time

And now it's time for some real action.

Continued next page

Day of Operation

This is the day when the plan finally meets reality. All that effort put into granting crew bids, placing smart buffers and tightly managing costs is now being put to use as crew fall ill or get delayed, aircraft have technical issues, and poor weather and congestion disrupt the operation causing compounding effects. The focus of the operation now shifts to moving passengers and cargo from A to B, almost any way possible. Only the regulatory flight and duty limits are sacred, but union rules can be discussed with the crew, and if needed, violated. Granting the bids and requests from crew is a nice-to-have on a busy day.

There are still many choices to make for reducing the fatigue risk, including choices around when to do stand-by call outs (potentially waking crew up), re-assigning production while taking advice from a BMM, etc. It may, for example, be that a flight is reinforced with an extra crew to compensate for a delay that made the duty more challenging.

The work in day of operation is often time-critical, which is why integrated and fast decision support is important. There is seldom time available for exporting rosters to an external system for a separate evaluation of fatigue risk. The functionality needs to be integrated within the crew tracking solution. Best practice includes using fatigue markers in the Gantt view, support for finding the most suitable standby crew, ability to sort/select/search while taking a BMM into account and producing detailed alertness graphs on demand.

It is important to stay local and minimalistic in the problem solving; it may be better to go with a solution to a problem that affects only three pilots, than a (seemingly) much cheaper one that affects ten. With detailed modeling of the operator's priorities around this trade-off and the costs involved, the system can automatically provide the crew tracker (a crew scheduler) with decision support.

Some operators have automatic alerts installed, providing a warning pushed to the crew tracker when crew are predicted to be below a certain threshold on alertness. Such warnings are not used as go/no-go decisions, but rather to trigger the crew tracker to contact the crew proactively to have a discussion around fitness for duty. This creates an extra awareness of risk for the crew, promoting prevention and mitigation actions.

The risk-increase in day of operation can be measured using AFR/NFR by establishing a measurement point just before, and another one just after, day of operation. Jeppesen Concert [4] is one such solution delivering this ability. By studying the worst cases (those flights that are moved into the concerning tail of the risk distribution) by this process step, insights can be gained that lead to improved guidance such as avoiding early departures the day after simulator duty.

Day of Operation Summary

- Take advice from BMM in roster updates
- Use a BMM to trigger consultation with crew
- Stay local in problem solving
- Quantify and track AFR/NFR change per fleet/rank/base over time

One important step of the process now remains.

Follow-up

In this step, following after the hectic day of operation, rosters are updated to correctly reflect what took place. Duty times and duty codes are corrected and various accounts are calculated and updated per crew such as number of landings, block time and duty time accumulators, airport recency, cosmic radiation accounts, etc. Also, typically monthly, the salaries and compensations are calculated from the roster content. It is made certain that safety/fatigue/occurrence reports are registered in relevant systems available for analysis.

Next follows the learning phase; what could have been done better in terms of controlling the fatigue risk exposure? Best practice includes analysis of the produced period in detail quantifying AFR/NFR looking for trends (ideally using a SPC approach, see Figure 5 and 3), but also going over the flights that predict the worst in more detail. This is done by producing alertness graphs and trying to find commonality leading up to insights on possible actions for enhancing control. Obviously, findings may relate to any of the preceding process steps being the cause of risk elevation. It is not certain that any suitable action regarding the roster production can be found that would drastically reduce risk. It may be that other controls for prevention and mitigation is all that remains, given the business model. A range of actions of this type for the organization and the crew to consider is found in supporting material [6].

So, what about fatigue reports? Fatigue reports have the advantage of complementing the analysis above with facts around instances of actual experienced sleepiness. For that reason, they are extremely valuable to analysis and should be actioned similarly to other BMM-driven findings. However, best practice is to also recognize that there are significant differences between individuals in how, and to what extent, they experience fatigue. Crew are different in how they prioritize, prepare and succeed in obtaining sleep when opportunities exist. This, in combination with the fact that the organization is typically receiving a very low number of fatigue reports relative to the total amount of flight assignments, makes fatigue reports likely to primarily reflect outliers in the data rather than assessing the real structure of risk exposure. If there is only data on 0.1% of the operation, are we really working data-driven? And what about potential adverse effects of the actions we are considering?

It is important to at least remove 'reasonable doubt' around the representativeness of the data before relying on findings. Best practice here would be to include regular and deeper data collection on routes or rostering patterns where there are indications of issues. (Also healthy individuals with a good night's sleep can now and then experience elevated sleepiness.)

Ideally, an operator would collect information on usage of controlled rest as well as self-assessments of sleepiness or fatigue at top of descent, on all flights performed. This type of collection will not deliver a complete picture either, but self-assessments are what drive fatigue reporting and complaints.

Such a collection is easy and inexpensive to operationalize if an operator desires a more detailed picture of the state of the operation. Operators with this type of data can reach deeper insights around the variability between crew operating the same patterns, and can direct their attention and actions more precisely to the part of the operation likely to deliver the best return of investment.

Once a finding is confirmed, it is easy to jump into action too quickly with a fix without investigating potential adverse effects. If long duties overnight are an issue for fatigue risk, is it really the case that shortening them will reduce the risk? In other words, does it reduce the overall fatigue risk, which is what should be addressed? What if the result is a much higher number of night duties - is that necessarily better? Each duty will have commute time before and after, and due to the higher number, there is now a need to plan them consecutively more often.

Best practice includes quantifying fatigue risk, for example, using AFR/NFR, and verifying an improvement on the overall situation, not only single flight duties in isolation. With good metrics in place, most operators already have reporting functionality enabling them to look at the full operation over a planning period and quantify the effect. Without quantification, there is a risk of spending time and effort increasing the fatigue risk exposure for the operation.

Follow-up Summary

- Quantify and track AFR/NFR change per fleet/rank/base over time
- Perform analysis of fatigue reports
- Consider deeper targeted data collection
- Consider implementing mandatory collection of self assessments (KSS or SP) at TOD
- Take caution when going from finding to fixing; verify
- Propose preventing and mitigating actions

Conclusion

The considerations listed in this document focus almost entirely on the crew management process and how to best sequence crew activities on the rosters through planning and operations. There are, of course, a lot of other actions to consider such as choice of hotel, SOPs for controlled rest, arranging other sleep facilities on ground, etc. Equally, there are a lot of different actions crew can take for best coping with the production they are given. Please find more written about those actions in the referenced material below [6].

Finally, there are plenty of other metrics than the alertness perspective coming from a BMM, which potentially can be used for quantifying the context of a flight. Such metrics can provide insights about flights being planned or flown in unusual ways compared to others. Please see [5, 7] below for more information and examples.

Additional reading:

[1] A Best Practice for Quantifying Fatigue Risk

[2] Aligning Rules With Human Physiology

[3] Are your processes in control?

[4] Jeppesen Concert

[5] Assignment-centric Performance Indicators

[6] Fatigue Risk Prevention and Mitigation

[7] BAM Safety Performance Indicators

[8] The secret behind pro-active risk reduction

[9] Your opinion is interesting, but...

Learn more about what we offer.

Find out how our Fatigue Risk Management System can optimize your operations by contacting your account representative or visiting jeppesen.com/frm.