USAFSAM Aeromedical Consultation Service Medical Risk Assessment and Airworthiness Matrix

Ryan S. Mayes; Christopher J. Keirns; Amy G. Hicks; Luke D. Menner; Maximilian S. Lee; Joseph H. Wagner; Robert L. Baltzer

INTRODUCTION: The 1% rule has long been a standard threshold for aerospace medical risk acceptance, but medical literature has noted multiple shortcomings with this threshold. Previous studies have suggested a risk matrix approach in aeromedical decision-making. General use of risk matrices for risk assessment is already codified in the U.S. Air Force (USAF). Based on this, the USAF School of Aerospace Medicine (USAFSAM) Aeromedical Consultation Service (ACS) generated and evaluated the ACS Medical Risk Assessment and Airworthiness Matrix (AMRAAM).

METHODS: The ACS adapted existing USAF standards to build the AMRAAM, gathered expert feedback, and sampled 100 previously adjudicated cases to compare legacy case dispositions to AMRAAM dispositions using polychoric correlation.

RESULTS: The AMRAAM disposition showed strong agreement with legacy dispositions ($\rho^* = 0.9424$). One case was discarded as it did not meet inclusion criteria. Of the 99 remaining cases, 88 had perfect agreement between legacy and AMRAAM dispositions. With the AMRAAM, eight cases were less restrictive and three were more restrictive (two due to an erroneous omission in the legacy disposition).

DISCUSSION: The AMRAAM produces disposition recommendations that are highly consistent with the legacy approach informed by the 1% rule, with discordant AMRAAM dispositions tending to be more permissive. The USAFSAM AMRAAM allows a more dimensional risk evaluation than the 1% rule, communicates aeromedical risk consistent with nonmedical USAF organizations, and harmonizes aeromedical risk with the level of risk the USAF has defined for all flying systems. The ACS will use the AMRAAM as standard practice in future aeromedical risk assessments.

KEYWORDS: aeromedical risk, aerospace medicine, risk assessment matrix, 1% rule, airworthiness.

Mayes RS, Keirns CJ, Hicks AG, Menner LD, Lee MS, Wagner JH, Baltzer RL. *USAFSAM Aeromedical Consultation Service Medical Risk Assessment and Airworthiness Matrix*. Aerosp Med Hum Perform. 2023; 94(7):514–522.

The 1% rule has long been a standard threshold for aerospace medical risk acceptance. The theoretical framework for the 1% rule began in British and European cardiology workshops in the 1980s[15,16] and subsequently has become the most widely accepted standard for aeromedical risk tolerance.[1,2,10] The 1% rule is the threshold of choice for the International Civil Aviation Organization;[7] while the U.S. Federal Aviation Administration does not explicitly refer to the 1% rule in its guidelines, it is generally concordant with international standards in defining high risk.[4] The 1% rule was developed for civilian aviation and targeted an all-cause fatal mishap rate of no more than 0.1 per million flight hours; the context for the calculations was dual-piloted commercial operations. The developers estimated that crew failures should account for no more than 10% of all fatal mishaps, and that no more than 10% of these crew failures should be due to underlying medical conditions causing incapacitation. It was further estimated that 1 in 1000 such incapacitations would occur in a situation in which the second pilot would be unable to recover the aircraft; this came from an estimation that only 10% of the average 1-h flight

From the U.S. Air Force School of Aerospace Medicine, Wright-Patterson AFB, OH, USA.

This manuscript was received for review in July 2022. It was accepted for publication in March 2023.

Address correspondence to: Ryan Mayes, Ph.D., M.P.H., Department of Aerospace Medicine, U.S. Air Force School of Aerospace Medicine, 2510 Fifth Street, W116F, Wright-Patterson AFB, OH 45433, USA; ryan.mayes.2@us.af.mil.

DOI: https://doi.org/10.3357/AMHP.6154.2023

time (in the early 1980s) would be considered critical, and that only one out of every 100 events would occur in conditions where the other pilot was unable to take control. Given these considerations, the developers calculated that an acceptable risk of medical incapacitation was one incapacitation per million flight hours. To annualize the risk, they further assumed 10,000 h/yr (simplified from 8760 h/yr); 1,000,000 h divided by 10,000 h/yr yielded an incapacitation rate of 1 in 100 per year, or 1%.[2,15,16] This simple rule of thumb provided a relatable reference point when making aeromedical decisions.

Despite its widespread use, the 1% rule is not without controversy. Mitchell and Evans noted that the 1% rule only accounted for total incapacitation, while most pilot incapacitations came from noncardiovascular causes and would not necessarily cause complete incapacitation.[10] Mitchell and Evans also advocated for a recalculation of the 1% rule accounting for an average flight time of 2 h, reasoning that longer flights would have a lower proportion of critical phases than was originally calculated. Finally, Mitchell and Evans observed that acceptable engine failure risk in twin-engine aircraft was approximately 5.8% per year, which was higher than the acceptable levels of aircrew incapacitation. Evans notes that, while the 1% rule provides a useful objective standard, there are many potential concerns with its use; for example, that cardiovascular mortality rates do not necessarily correspond with incapacitation rates, that cardiovascular incapacitation is not only caused by heart attacks, and that incapacitation does not occur due to cardiovascular disease alone.[3] Multiple manuscripts argue that age should be incorporated into aeromedical risk evaluation;[2,4,10] this is especially the case as cardiovascular risk is closely associated with age[11] (as was acknowledged by the developers of the 1% rule in their 1988 paper[15]). Given these concerns with the 1% rule, Evans argues that it may be too restrictive.[2,3] Despite noted issues with the 1% rule, it remains a commonly cited threshold in aeromedical risk assessment for cardiac and noncardiac conditions causing incapacitation, including within the U.S. Air Force (USAF).

The USAF School of Aerospace Medicine (USAFSAM) is the USAF organization focused on education, operational consultation, and research in aerospace and operational medicine. Established in 1918, USAFSAM is one of the oldest and largest continually operating military flight medicine centers of expertise in the world. The Aeromedical Consultation Service (ACS) within USAFSAM has provided expert aeromedical consultation to the USAF since the 1950 s. The ACS evaluates USAF aircrew with medical conditions that disqualify them from their duties; these aircrew require waivers to continue flying duties. The ACS analyzes these waiver requests and produces over 2000 individualized aeromedical risk assessments per year through in-house expertise in aerospace medicine and other specialties. These risk assessments are used by USAF decision authorities in waiver disposition.

The ACS has used the 1% rule as a basis for aeromedical disposition recommendations for incapacitating medical events since the aeromedical community embraced this threshold, but has also recognized that the 1% rule was insufficient to determine risk for non-incapacitating medical events and may not accurately reflect the risk tolerance of the USAF operational community (the Line of the Air Force, or LAF). In 2017 and 2018, the USAF sought to increase the pool of medically qualified USAF pilot applicants for an expanded training pipeline, resulting in increased scrutiny of all medically disqualified applicants. This renewed interest within the ACS to formalize quantitative risk thresholds that went beyond the 1% rule, as well as a desire to improve the quality of communication and clarify the rationale for aeromedical disposition recommendations to the LAF. The ACS was aware of previous publications using risk matrices to conceptualize aeromedical risk,[5,6] and identified a risk matrix tool as the most viable solution to standardize a quantitative threshold and improve communication. The LAF already uses a risk matrix approach to assess risk in other domains, prompting the ACS to begin development of an aeromedical risk matrix in early 2020.

The LAF use of a risk matrix approach to assess risk is driven by policy. Air Force Instruction (AFI) 90-802 directs the use of risk management principles across the USAF and defines procedural steps in generating an overall risk assessment.[18] Separate assessments of hazard probability and severity are foundational to the formation of a risk matrix. Risk matrices are further defined in Air Force Pamphlet (AFPAM) 90-803; **Fig. 1** provides a sample risk matrix from that publication.[19]

USAF Airworthiness Bulletin (AWB) 150B echoes the approach of separately identifying severity and probability, and using a risk matrix to determine an overall risk level.[17] AWB-150B establishes specific definitions for four severity categories and six probability ranges (though one probability category is zero, through elimination of risk), and integrates them into four overall risk levels using a risk matrix (**Fig. 2**): high, serious, medium, and low. AWB-150B enacts AFI 62-601, which defines airworthiness as the "property of an air system configuration to safely attain, sustain, and terminate flight."[21] Taken together, AFI 62-601 and AWB-150B effectively establish risk criteria for air systems and components of those systems. In turn, AFI 62-601 provides direction based on policy established by Air Force Policy Document 62-6, which specifies that the purpose of airworthiness is to provide USAF personnel an appropriate level of safety of flight, establishing a linkage between weapons systems and their operators.[22] The recently updated version of AFPAM 90-803 continues this line of reasoning, stating "To apply the systematic [risk management] process, the composition of hardware, procedures, and people that accomplish the mission or produce mishaps, should be viewed as a system."[20]

The need for more granular risk assessments is not novel in aerospace medicine. As noted above, multiple authors have advocated for updates or modifications to the 1% rule. The concept of risk as the product of the likelihood and severity of an adverse aeromedical event has precedent as well.[8,13] Prudhomme et al. used this construct to evaluate the overall risk of multiple pharmaceuticals, but noted challenges with identifying a level of acceptable risk.[13] Gray, Sargsyan, and Davis proposed a risk matrix approach for establishing acceptable clinical risk levels for long duration space missions in 2010.[6] This is the earliest

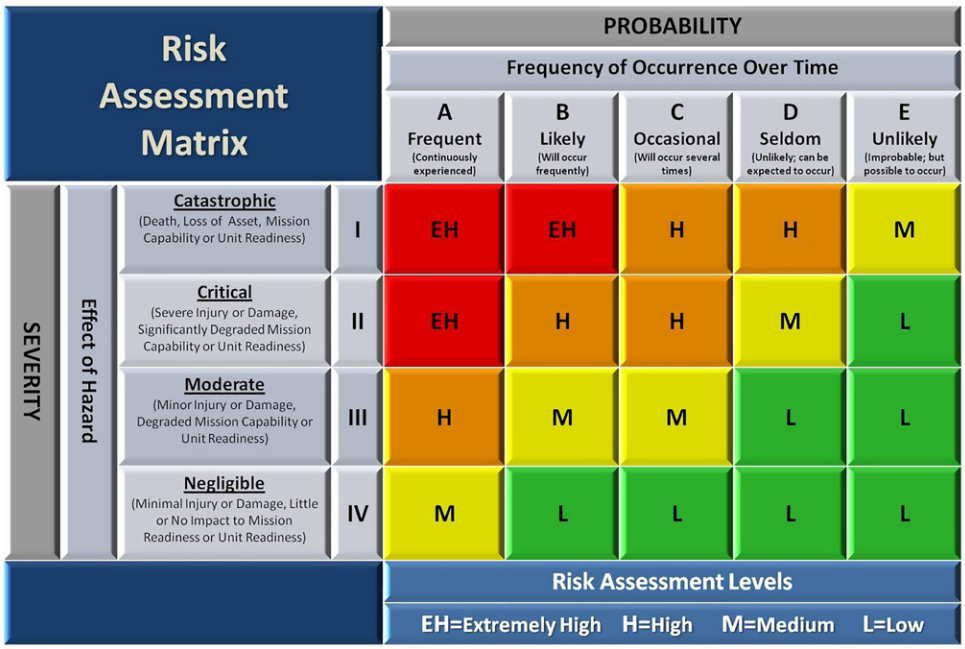


Fig. 1. Sample risk assessment matrix from AFPAM 90-803.[19]

example of a published aeromedical risk matrix found by a comprehensive literature review, and did include levels of acceptable risk. The authors found that the risk matrix approach promoted evidence-based decision making and was broadly applicable.[6] This approach was extended by Gray et al. in the development of three-dimensional risk matrices to assess aeromedical risk.[5] As with the 1% rule, these matrices were based on a cardiology working group, and incorporated likelihood and severity, with the third dimension being crew position. This three-dimensional approach helped inform the present study. USAFSAM developed and tested the Aeromedical Consultation Service Medical Risk Assessment and Airworthiness Matrix (AMRAAM) to address the limitations of the 1% rule, to align aeromedical risk analyses with USAF guidance, and to better communicate aeromedical risk to the USAF operational community. The similarity in name to the missile platform is deliberate and reflects the operational relevance upon which the matrix was built.

METHODS

Development of the USAFSAM AMRAAM

The USAFSAM ACS had previously identified a need to better communicate aeromedical risk, as well as a need to determine whether aeromedical risk thresholds were appropriate and

USAF Airworthiness Risk Assessment Matrix			Severity Category			
Probability Level	Probability per FH or Sortie	Freq per 100K FH or 100K Sorties	Catastrophic (1)	Critical (2)	Marginal (3)	Negligible (4)
Frequent (A)	$10^{-3} \leq$ Prob	100 ≤ Freq	1	3	7	13
Probable (B)	$10^{-4} \leq$ Prob < 10^{-3}	10 ≤ Freq < 100	2	5	9	16
Occasional (C)	$10^{-5} \leq$ Prob < 10^{-4}	1 ≤ Freq < 10	4	6	11	18
Remote (D)	$10^{-6} \leq$ Prob < 10^{-5}	0.1 ≤ Freq < 1	8	10	14	19
Improbable (E)	0 < Prob < 10^{-6}	0 < Freq < 0.1	12	15	17	20
Eliminated (F)	Prob = 0	Freq = 0	Eliminated			

High	RAC = 1 - 5	Medium	RAC = 10 – 17
Serious	RAC = 6 – 9	Low	RAC = 18 – 20

Fig. 2. U.S. Air Force (USAF) airworthiness risk assessment matrix.[21] The USAF Airworthiness Bulletin defines specific probability/frequency levels, as well as descriptions for each of four severity categories. The product of probability/frequency yields a risk; these risks are grouped into one of four risk assessment codes (RACs).

consistent with broader USAF risk acceptance. The use of risk matrices offered a solution to both issues. In order to communicate effectively and to clearly connect to existing USAF risk acceptance thresholds, the USAFSAM AMRAAM was derived directly from AWB-150B, which establishes levels of acceptable risk for air systems. Because Air Force Policy Document 62-6 explicitly links system safety to human safety, the AWB definitions were chosen to ensure that aeromedical risk was aligned to existing USAF risk acceptance thresholds. As with the 1% rule, this approach harmonizes aeromedical risk to a nonmedical threshold. Unlike the 1% rule, however, the AMRAAM harmonizes aeromedical risk to an existing and accepted standard.

The AMRAAM began with "probability" by examining and translating the definitions found in the AWB. Of note, even though the AFPAM and AWB-150B use the term "probability," the AMRAAM uses the term "likelihood" to express the projected chance of a medical event. While the terms confidence, likelihood, and probability are often used interchangeably on an informal basis, the term likelihood is preferred in the AMRAAM as it is more technically accurate—likelihood connects both known and uncertain data.[14] This terminology is consistent with Gray et al.[5] The AMRAAM uses the same five likelihood categories and titles provided by the AWB (frequent, probable, occasional, remote, and improbable). However, these definitions required adjustments in order to better reflect likelihood data commonly reported in medical literature; while the AWB expresses likelihood in terms of flight hours or sorties, most medical literature provides denominators in years or person-years. To bridge this gap, the ACS calculated equivalent likelihoods for each category by annualizing the risk thresholds defined by the AWB. For instance, an "occasional" event would occur once every 10,000 (10^4) to 100,000 (10^5) h. Over the course of a calendar year, the likelihood of one event occurring within these defined parameters would range from 8.39 to 58.38% using the following calculations:

$$1-\left[1-\left(\frac{1}{10^5}\right)\right]^{365.25*24}=0.0839$$

$$1-\left[1-\left(\frac{1}{10^4}\right)\right]^{365.25*24}=0.5838$$

In general, the AMRAAM rounds up, which accepts slightly more risk than is represented in the AWB; the AMRAAM defines the likelihood of a single occurrence per year for the "occasional" category as 10–60%. The AMRAAM also displays these annualized likelihoods in equivalent 5- and 10-yr timeframes in order to assist in translation between medical literature and the AMRAAM.

The severity categories from the AWB (catastrophic, critical, marginal, negligible) were also carried forward into the AMRAAM. The definitions for each category were informed by AWB-150B, AFPAM 90-803, and AFI 90-802. This allowed the severity categories to account for mission impact, flying safety, crew position, specific airframe, and aircrew health, while still corresponding to the AWB framework.

The development process resulted in a completed draft version of the AMRAAM; this approach allows providers to assess the likelihood of an aeromedical event of concern by selecting the appropriate column and to assess the impact to mission, flight safety, and aircrew health by selecting the appropriate row. This provides a specific risk score and risk assessment level, which corresponds directly to the AWB.

This initial version of the AMRAAM appeared to be an effective communication tool and a framework in conceptualizing baseline and mitigated risk. To evaluate both of these aspects, the study team proceeded to gather subject matter expert inputs on the AMRAAM.

Qualitative Assessment of the USAFSAM AMRAAM

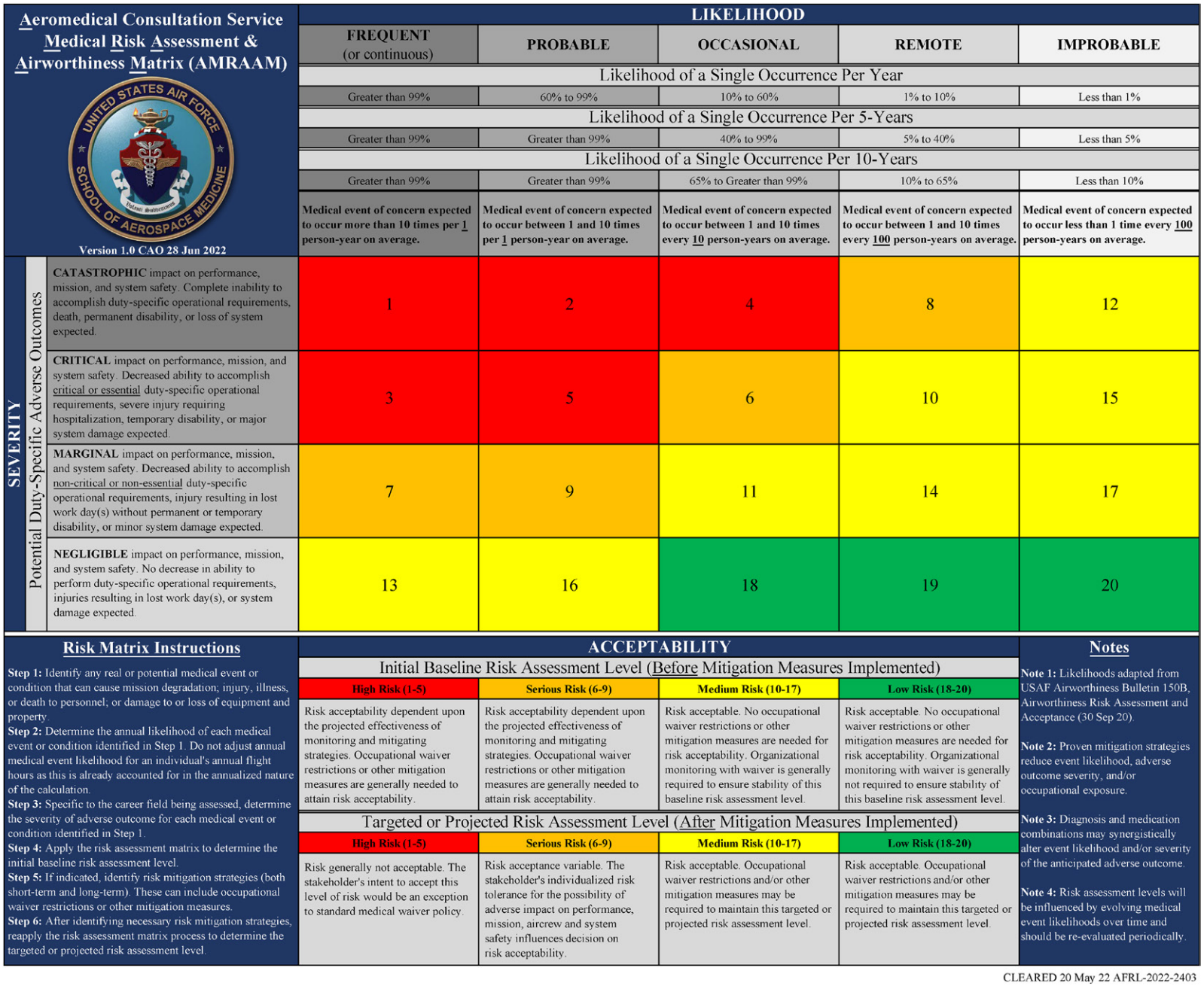
The initial draft of the AMRAAM was used by ACS flight surgeons in a tabletop simulated review of 50 ACS cases. This exercise established the feasibility of the risk matrix approach and potential applicability of an airworthiness standard for USAF aeromedical risk assessment.

Draft versions of the AMRAAM were iteratively reviewed with experienced flight surgeons and senior aerospace medicine specialists across the USAF, which resulted in refinements to the presentation of likelihood and severity, as well as to the instructions for use. The draft was also reviewed with LAF members of the Air Force Safety Center to gather inputs on whether the AMRAAM was consistent with flight safety practices. Based on inputs from these subject matter experts, the study team produced "Version 1.0" of the AMRAAM (**Fig. 3**). Version 1.0 was used in the next phase of the study—an initial statistical validation of the AMRAAM.

Statistical Assessment of the USAFSAM AMRAAM

Subjects. No new data were collected for this study. Subjects consisted of 100 randomly selected cases evaluated by the ACS from 1 January 2019 through 31 December 2019; this time period was selected because it occurred prior to development of the AMRAAM while still aligning with current USAF aeromedical policy. Because the ACS performs both in-person evaluations and conducts remote reviews, the 100 total cases included 50 of each category. The test statistic was polychoric correlation, but a sample size calculation for this statistic was not available. However, a sample size calculation for a one-sample correlation test suggested that $N = 100$ would be sufficient to detect a minimum difference in correlation of approximately 0.27 or less (this minimum difference decreases as the a priori correlation increases). Because it was anticipated that there would be relatively high correlation between legacy and AMRAAM dispositions (0.6 or higher), a sample size of 100 was anticipated to be sufficient with alpha = 0.05 and a power of 0.80.

Inclusion criteria specified cases would be Flying Class II (manned aircraft) pilots, each case had previously been reviewed by the ACS, and each case had a completed disposition recommendation, with one of four possibilities: medically qualified, unrestricted waiver, restricted waiver, or disqualified. Remotely piloted aircraft pilots, navigators, and flight surgeons were excluded from the study. In addition, cases were excluded if the



Aeromedical Consultation Service Medical Risk Assessment & Airworthiness Matrix (AMRAAM)

Version 1.0 CAO 28 Jun 2022

SEVERITY: Potential Duty-Specific Adverse Outcomes	LIKELIHOOD: FREQUENT (or continuous)	PROBABLE	OCCASIONAL	REMOTE	IMPROBABLE
	Likelihood of a Single Occurrence Per Year				
	Greater than 99%	60% to 99%	10% to 60%	1% to 10%	Less than 1%
	Likelihood of a Single Occurrence Per 5-Years				
	Greater than 99%	Greater than 99%	40% to 99%	5% to 40%	Less than 5%
	Likelihood of a Single Occurrence Per 10-Years				
	Greater than 99%	Greater than 99%	65% to Greater than 99%	10% to 65%	Less than 10%
	Medical event of concern expected to occur more than 10 times per 1 person-year on average.	**Medical event of concern expected to occur between 1 and 10 times per 1 person-year on average.**	**Medical event of concern expected to occur between 1 and 10 times every 10 person-years on average.**	**Medical event of concern expected to occur between 1 and 10 times every 100 person-years on average.**	**Medical event of concern expected to occur less than 1 time every 100 person-years on average.**
CATASTROPHIC impact on performance, mission, and system safety. Complete inability to accomplish duty-specific operational requirements, death, permanent disability, or loss of system expected.	1	2	4	8	12
CRITICAL impact on performance, mission, and system safety. Decreased ability to accomplish critical or essential duty-specific operational requirements, severe injury requiring hospitalization, temporary disability, or major system damage expected.	3	5	6	10	15
MARGINAL impact on performance, mission, and system safety. Decreased ability to accomplish non-critical or non-essential duty-specific operational requirements, injury resulting in lost work day(s) without permanent or temporary disability, or minor system damage expected.	7	9	11	14	17
NEGLIGIBLE impact on performance, mission, and system safety. No decrease in ability to perform duty-specific operational requirements, injuries resulting in lost work day(s), or system damage expected.	13	16	18	19	20

Risk Matrix Instructions

Step 1: Identify any real or potential medical event or condition that can cause mission degradation; injury, illness, or death to personnel; or damage to or loss of equipment and property.
Step 2: Determine the annual likelihood of each medical event or condition identified in Step 1. Do not adjust annual medical event likelihood for an individual's annual flight hours as this is already accounted for in the annualized nature of the calculation.
Step 3: Specific to the career field being assessed, determine the severity of adverse outcome for each medical event or condition identified in Step 1.
Step 4: Apply the risk assessment matrix to determine the initial baseline risk assessment level.
Step 5: If indicated, identify risk mitigation strategies (both short-term and long-term). These can include occupational waiver restrictions or other mitigation measures.
Step 6: After identifying necessary risk mitigation strategies, reapply the risk assessment matrix process to determine the targeted or projected risk assessment level.

ACCEPTABILITY

Initial Baseline Risk Assessment Level (Before Mitigation Measures Implemented)

High Risk (1-5)	Serious Risk (6-9)	Medium Risk (10-17)	Low Risk (18-20)
Risk acceptability dependent upon the projected effectiveness of monitoring and mitigating strategies. Occupational waiver restrictions or other mitigation measures are generally needed to attain risk acceptability.	Risk acceptability dependent upon the projected effectiveness of monitoring and mitigating strategies. Occupational waiver restrictions or other mitigation measures are generally needed to attain risk acceptability.	Risk acceptable. No occupational waiver restrictions or other mitigation measures are needed for risk acceptability. Organizational monitoring with waiver is generally required to ensure stability of this baseline risk assessment level.	Risk acceptable. No occupational waiver restrictions or other mitigation measures are needed for risk acceptability. Organizational monitoring with waiver is generally not required to ensure stability of this baseline risk assessment level.

Targeted or Projected Risk Assessment Level (After Mitigation Measures Implemented)

High Risk (1-5)	Serious Risk (6-9)	Medium Risk (10-17)	Low Risk (18-20)
Risk generally not acceptable. The stakeholder's intent to accept this level of risk would be an exception to standard medical waiver policy.	Risk acceptance variable. The stakeholder's individualized risk tolerance for the possibility of adverse impact on performance, mission, aircrew and system safety influences decision on risk acceptability.	Risk acceptable. Occupational waiver restrictions and/or other mitigation measures may be required to maintain this targeted or projected risk assessment level.	Risk acceptable. Occupational waiver restrictions and/or other mitigation measures may be required to maintain this targeted or projected risk assessment level.

Notes

Note 1: Likelihoods adapted from USAF Airworthiness Bulletin 150B, Airworthiness Risk Assessment and Acceptance (30 Sep 20).

Note 2: Proven mitigation strategies reduce event likelihood, adverse outcome severity, and/or occupational exposure.

Note 3: Diagnosis and medication combinations may synergistically alter event likelihood and/or severity of the anticipated adverse outcome.

Note 4: Risk assessment levels will be influenced by evolving medical event likelihoods over time and should be re-evaluated periodically.

CLEARED 20 May 22 AFRL-2022-2403

Fig. 3. USAFSAM Aeromedical Consultation Service Medical Risk Assessment and Airworthiness Matrix (AMRAAM). Version 1.0 of the USAFSAM AMRAAM was used in the initial validation testing.

case was returned to a requesting entity to collect more information, the recommendation was to continue a temporary "Duties Not Including Flying" status, or if the disposition did not have a final recommendation. The Air Force Research Laboratory Institutional Review Board determined this project did not meet the regulatory definition of human subject research, documented as protocol FWR20210154N, 9 June 2021.

Procedure. One member of the investigatory team was responsible for reviewing and sanitizing those cases meeting inclusion criteria by removing any individual identifiers; this member did not participate in case reviews for the validation portion of the study. Per the inclusion criteria, each case already had an existing disposition recommendation which was considered the "legacy" recommendation for the study. Each de-identified case was evaluated using the AMRAAM by at least one physician within the specialty appropriate for the case; specialties included cardiology, internal medicine, neurology, ophthalmology, psychiatry, pulmonology, and sleep medicine. In order to replicate the administrative process by which the legacy recommendations were generated, cases that were originally seen in person were presented and discussed in a case conference led by an aerospace medicine specialist to determine a disposition recommendation for each case. This conference included participation from all specialties. Cases that were initially record reviews were reviewed by the relevant specialty or specialties, followed by a meeting with at least one aerospace medicine specialist to determine the disposition recommendations. These processes mirrored the deliberative case workflows used by the ACS.

For each case, the clinical specialists were instructed to identify all aeromedical events of concern and separately select the corresponding likelihood category and severity category for each aeromedical event. The combination of likelihood and severity identified one specific cell in the AMRAAM per aeromedical event, generating a risk score as shown in Fig. 3. Each specialist provided their individual scores in a review with aeromedical and operational specialists. Through a collaborative process, a consensus risk score was assigned for each aeromedical event; the highest of these scores was used to determine the risk assessment level designated in the AMRAAM (low, medium, serious, or high).

The next step was determined by the risk assessment level. If that level was low or medium, no occupational restrictions or other mitigation measures were needed to attain risk acceptability. If the initial risk assessment level was serious or high, this prompted the reviewers to consider occupational restrictions or other mitigation measures to either reduce the likelihood category, severity category, or both. If mitigation strategies were appropriate, a new risk score accounting for these mitigations was calculated to determine the targeted risk assessment level. This level was used to inform overall risk acceptability and the final aeromedical disposition recommendation for the case. The possible disposition recommendations mirrored those from the legacy approach: medically qualified, unrestricted waiver, restricted waiver with specified limitations, or disqualified. The AMRAAM disposition recommendation was determined solely for the purpose of evaluating the AMRAAM and was not shared with entities external to the ACS.

Once each case received an AMRAAM disposition, the results were collated to produce a comparison of the legacy recommendation to the AMRAAM recommendation. ACS providers then discussed any discrepancies between legacy and AMRAAM recommendations to identify potential reasons for differing dispositions.

Statistical analysis. The design of the initial validation study provided a direct comparison of the legacy vs. AMRAAM disposition recommendation: each case received a disposition of medically qualified, unrestricted waiver, restricted waiver, or disqualified. These dispositions are categorical and are ordinal (listed from least restrictive to most restrictive). The AMRAAM case review process was designed to mirror the process used for virtual and in-person case reviews; if the legacy and AMRAAM processes led to similar dispositions, it could be assumed that the AMRAAM did not substantially impact the overall assessment. Conversely, a lack of association between legacy and AMRAAM dispositions would mean that the two processes yielded different results based on the same data, which may have implications for the validity of one or both processes. Because it was expected that the legacy and AMRAAM dispositions would be related, and because the outcomes for both were ordinal, polychoric correlation was selected. Polychoric correlation is a technique for estimating the correlation between two observed ordinal variables. The correlation coefficient is represented as ρ^*, ranges from 0 to 1, and is interpreted in the same way that a Pearson's correlation coefficient would be, with 0 indicating no relationship between variables and 1 representing a perfect correlation.[12] Under this test, the null hypothesis is that there is no relationship between the legacy and AMRAAM dispositions.

RESULTS

As discussed in the Methods section, the qualitative analysis of the AMRAAM resulted in the final product shown in Fig. 3. After the sample of 100 cases was generated and analysis began, one case was excluded because it did not meet inclusion criteria; specifically, it did not have a final disposition recommendation. This left 99 cases that met all inclusion criteria. Case disposition recommendations from the AMRAAM vs. legacy recommendations are shown in **Fig. 4**. The AMRAAM disposition showed strong agreement with legacy dispositions, with $\rho^* = 0.9424$ ($P \ll 0.0001$). This association is highly statistically significant and indicates that the AMRAAM produces results that are strongly correlated with the legacy process; this level of correlation is regarded as "almost perfect" correlation.[9] Of note, this analysis included two cases that had different outcomes due to a change in aeromedical policy; had those cases been omitted the correlation coefficient would be even higher.

Of the 99 cases, 90 had the same overall disposition recommendation from the legacy and AMRAAM process and 88 had

		AMRAAM Disposition				
		Qualified	Unrestricted Waiver	Restricted Waiver	Disqualified	Total
Legacy Disposition	Qualified	3 (100%)	0	0	0	3
	Unrestricted Waiver	2 * (2.8%)	67 (94.4%)	1 † (1.4%)	1 † (1.4%)	71
	Restricted Waiver	0	5 (25%)	15 ‡ (75%)	0	20
	Disqualified	0	0	0	5 (100%)	5
Total		5 (5.1%)	72 (72.7%)	16 (16.2%)	6 (6.1%)	99

Fig. 4. Legacy disposition recommendation vs. AMRAAM disposition recommendation. The diagonal cells from top left to bottom right are concordant (white boxes), where the AMRAAM and legacy dispositions were the same. The AMRAAM disposition was more restrictive for cells above this diagonal line (medium gray), and less restrictive for cells below the line (light gray). Percentages are expressed as the cell number divided by the total legacy dispositions for a given row. *Both cases were impacted by a policy change. The AMRAAM and legacy dispositions were in accordance with aeromedical policy at the time of review; the policy changed in between legacy and AMRAAM dispositions. †The legacy disposition was not in accordance with aeromedical policy at the time of the legacy disposition recommendation. ‡Compared to the legacy disposition, one restricted waiver was less restrictive with the AMRAAM disposition, and one restricted waiver was more restrictive with the AMRAAM disposition.

exactly the same disposition; 2 cases had restricted waivers with both processes but had differing restrictions. Of the 11 total cases with some difference (either a different disposition or waivers with different restrictions), there was a clear tendency for the AMRAAM to be less restrictive than the legacy process (with 8 of 11 cases being less restrictive). Only three cases had a more restrictive disposition from the AMRAAM compared to the legacy process. Two of these cases reflected an erroneous omission in the legacy disposition; in each case the legacy process should have included an additional restriction. Neither of these dispositions affected flight safety during the period of the waiver but should have contained additional restrictions to prevent operational risk over the course of the aviator's career.

Of the 11 total cases with some difference (either different disposition or different restrictions), 8 had a less restrictive disposition from the AMRAAM compared to the legacy process. Two of these eight reflected a standards change—a less restrictive standard at the time of the AMRAAM study led to a qualified disposition, rather than the legacy unrestricted waiver disposition (which was appropriate in 2019). One of the eight was an AMRAAM restricted waiver with fewer restrictions than the legacy restricted waiver. The remaining five were cases in which the AMRAAM process resulted in unrestricted waivers rather than restricted waivers under the legacy process. One case would have resulted in an immediate operational impact for the current airframe. In five cases, the less restrictive AMRAAM recommendation expanded future career opportunities in other airframes that would not have been permitted under the legacy recommendation.

In total, only two cases where the AMRAAM and legacy recommendations differed would have made an immediate difference in the ability of a pilot to fly their current platform. One case where the AMRAAM yielded a less restrictive result would have allowed a fighter pilot to continue to fly in a high-performance airframe. A second case where the AMRAAM was more restrictive would have kept a nonhigh performance pilot out of the aircraft for another 1-2 mo to initiate therapy and monitor efficacy prior to waiver reconsideration. Additionally, the AMRAAM identified two recommendations in the legacy model which appear to have contained an erroneous omission; both omissions resulted in the legacy recommendation being an unrestricted waiver. The AMRAAM recommendations for both of these cases were more restrictive but were more consistent with aeromedical policy.

DISCUSSION

Construct and Application of the USAFSAM AMRAAM

The AMRAAM was developed to address the limitations of the 1% rule, to align aeromedical risk thresholds with USAF guidance, and to more effectively communicate aeromedical risk between USAF aeromedical and operational communities. There are at least two critical components that facilitate this improved communication: 1) alignment with risk assessment standards and practices already in use across the USAF; and 2) the ability to decompose risk into two dimensions (severity and likelihood). This added dimensionality in risk assessment allows aerospace medicine to be less of a "black box" when medical professionals communicate the reasoning for a risk determination to nonproviders. For comparison, the 1% rule is restricted to 1 of the 20 cells (cell 12) in the AMRAAM; the AMRAAM provides a much more detailed and dimensional analysis.

The dimensions specified in the AMRAAM provide an additional benefit. In complicated cases where multiple conditions require analysis, the AMRAAM supports a systematic approach to assessing likelihood and severity for each event of aeromedical concern, both before and after mitigating strategies are applied. This facilitates a cogent risk assessment that highlights the most important aeromedical aspects of these complex cases and the primary mitigating measures needed to reduce risk. Finally, the construct of the AMRAAM annualizes likelihood per flight hour, eliminating any need to factor total flight hours into the AMRAAM analysis.

In their 2019 paper, Gray et al. advocated for the use of risk matrices with a third dimension for occupational (aircrew) duty.[5] Development of the AMRAAM began with a similar concept. However, rather than developing separate matrices for different occupational roles, the AMRAAM incorporates the "third dimension" of aircrew position into the assessment of the severity of aeromedical events. This eliminates the need for separate matrices because the contribution of aircrew position is factored into the overall risk assessment through severity. For example, an aeromedical condition such as defective stereopsis may be high risk for a single-seat pilot, since the likelihood would be frequent and the severity would be critical or catastrophic. However, the overall risk can be reduced by a change in pilot duty such as flying a remotely piloted aircraft; this would not change the likelihood of occurrence but would reduce the

severity of impact. In addition, showing the effect of risk mitigation through operational restrictions on a single two-dimensional matrix simplifies communication and enables USAF medics to more clearly communicate the benefits of recommended risk mitigation strategies to the LAF. Because USAF assessments of aeromedical risk are effectively considering overall risk to a weapons system, the ability to show how risk may vary based on aircrew role may be more effective in two dimensions.

Finally, the AMRAAM presents an opportunity for more effective two-way communication with the LAF. The AMRAAM harmonizes risk assessment processes, language, and risk tolerance thresholds between the LAF and aeromedical communities. In doing so, the AMRAAM may serve as a catalyst for meaningful dialogue on acceptable levels of risk, mitigation strategies, and operational impact. Ultimately, this may enable a more transparent and collaborative decision-making process.

Initial Formal Validation of the USAFSAM AMRAAM

The AMRAAM shows clear benefits as a communication tool and as a platform to provide additional dimensionality in risk assessment. These benefits would be of little use if the AMRAAM was nonvalidated or unpredictable. However, a formal analysis shows that the AMRAAM produces results highly consistent with the ACS legacy approach; 88/99 cases reviewed had complete agreement between the AMRAAM and legacy processes. The polychoric correlation between the two processes is almost perfect,[9] and the statistical significance of the correlation coefficient verifies the relationship between AMRAAM and legacy dispositions. This effectively uses the legacy process as an accepted standard with which to evaluate the AMRAAM process; it is theoretically possible that the legacy process does not provide satisfactory results. However, because the ACS is an advisory body, it is possible to measure the quality of recommendations using the legacy ACS process, and >98% of ACS recommendations in 2019 were accepted by the waiver authority. Based on this, the legacy process appears to provide satisfactory results and is a reasonable benchmark for comparison. Of particular note, of the 11 nonconcordant cases for which the legacy and AMRAAM processes did not have complete agreement, the majority (8) were less restrictive with the AMRAAM disposition. Of the three cases that had a more restrictive AMRAAM disposition, two reflected errors in the legacy disposition. The nonconcordant cases are small in number but suggest that when the legacy and AMRAAM processes do differ, the AMRAAM tends to yield less restrictive recommendations. In addition, the AMRAAM identified two erroneous omissions from the legacy process; it is possible that AMRAAM results may be somewhat more reliable.

This study is not without limitations, with at least three areas of note: generalizability, bias due to recall or observation, and influence of policy changes. Regarding generalizability, the study involved application of the AMRAAM by clinically experienced ACS specialty consultants with operational aerospace medicine backgrounds. By virtue of its role, the ACS typically performs aeromedical risk assessment of complex, challenging cases. Therefore, while this random sample is likely valid for ACS cases, the study results may not apply to (or represent) all waivers for the USAF pilot population. Additionally, only Flying Class II (manned aircraft) pilot waivers were included in the study. Secondly, the legacy dispositions were made from cases received at the ACS from 1 January–31 December 2019. Although each case was de-identified, it is possible that the reviewers applying the AMRAAM could have remembered elements of some cases (or their dispositions), which could influence the outcome of this study. Observation bias is also a possibility; during the development of the AMRAAM, the risk matrix approach was refined with all members of the ACS through briefings and feedback sessions. The case reviewers' knowledge of the risk assessment tool and the potential tool adoption may have introduced observation bias if the reviewers assumed that the outcome of this study could facilitate organizational adaptation of a collaboratively developed tool. However, this potential limitation was mitigated by clinical information in the original case being de-identified and legacy recommendation outcomes being masked.

A third limitation is the potential for medical standards and policy changes since 2019 to influence the reassessment recommendations. Organizational culture, aeromedical adjudication experience, pilot personnel projections, newer published literature, therapeutic advancements, and internal/external tolerance to aeromedical risk constantly shape medical standards and policy, thereby influencing risk tolerance for various conditions. Two cases in this study were impacted by a policy change which affected aeromedical dispositions. This discrepancy was easily identified and no other policy changes impacting aeromedical disposition for cases in this study occurred between 2019 and the time of the study.

Beyond the validation study, there is one important limitation of the AMRAAM itself. Gray, Sargsyan, and Davis argued in 2010 that risk matrices had the potential to project more objectivity than was actually present and argued that the risk matrix approach was best used as a means to facilitate discussion, rather than as a risk-making tool.[6] The context for the AMRAAM is quite different. There are existing USAF standards for acceptable levels of risk in the system and the AMRAAM allows USAFSAM to easily harmonize to those existing standards. This, in turn, allows medical considerations to avoid adding excessive risk to an overall weapons system and to avoid being overly conservative, which may affect mission readiness. Gray et al. propose that risk matrices may mask uncertainty stemming from low levels of evidence; while this proposal is certainly valid, the USAF operational mission often dictates that decisions be made based on the best available evidence at the time. In the opinion of the authors, the ability to harmonize with overall USAF system risk far outweighs the potential drawbacks; we believe that basing aeromedical decision on the USAFSAM AMRAAM is the best available approach to assessing and communicating aeromedical risk in the USAF.

Finally, there is one important consideration regarding use of the AMRAAM for non-USAF applications. While the study team believes the construct of the AMRAAM is robust, it is important to highlight that the likelihood and severity scales

were directly informed by existing USAF standards. Other services, entities, or nations may not have the same level of risk tolerance. Caution should be exercised in any potential application of the AMRAAM to aeromedical assessments outside of the U.S. Air Force. Of note, the construct of the AMRAAM allows risk acceptability (indicated at the bottom of Fig. 3) to be tailored by adjusting the range of risk scores included in each risk assessment level. This would allow organizations to tailor the matrix to their specific risk tolerance.

Conclusion

The USAFSAM AMRAAM is a new paradigm for USAF aeromedical risk assessment. By decomposing risk into likelihood and severity, the AMRAAM allows a more dimensional analysis than does the 1% rule. This dimensionality also facilitates a more systematic approach to risk assessment by providing an objective process and clear definitions for risk, likelihood, and severity. The separation of likelihood and severity reflects the overall construct of the AMRAAM, which integrates LAF risk management processes and airworthiness standards. This allows the adoption of stakeholder risk tolerance into aeromedical risk assessment and promotes a human systems integration approach for aerospace medicine. These attributes enable USAF medics to more effectively communicate aeromedical considerations to the LAF, and may facilitate a more meaningful dialogue on operational risks and mitigation strategies. The USAFSAM Aeromedical Consultation Service will use the AMRAAM as the new basis for aeromedical risk assessments; the AMRAAM allows the ACS to harmonize aeromedical risk with USAF airworthiness standards, optimizing human performance for the overall air system.

ACKNOWLEDGMENTS

The views expressed are those of the authors and do not reflect the official guidance or position of the U.S. Government, the Department of Defense (DoD), or the U.S. Air Force. The appearance of external hyperlinks does not constitute endorsement by the DoD of the linked websites, or the information, products, or services contained therein. The DoD does not exercise any editorial, security, or other control over the information you may find at these locations.

Financial Disclosure Statement: The authors have no competing interests to disclose.

Authors and Affiliations: Ryan S. Mayes, Ph.D., M.P.H., Christopher J. Keirns, D.O., Amy G. Hicks, M.D., Luke D. Menner, D.O., Maximilian S. Lee, M.D., M.P.H., and Robert L. Baltzer, M.D., U.S. Air Force School of Aerospace Medicine, Wright-Patterson AFB, OH, USA; and Joseph H. Wagner, M.P.H., JYG Innovations, LLC, Dayton, OH, USA.

REFERENCES

1. Ernsting J, Nicholson AN, Rainford DJ, editors. Aviation medicine. 3rd ed. Waltham (MA): Butterworth-Heinemann; 1999:223–227.
2. Evans A. Aeromedical risk: A numerical approach. In: Rainford D, Gradwell D, editors. Ernsting's aviation medicine. 4th ed. London: Hodder-Arnold; 2006:223–227.
3. Evans A. International regulation of medical standards/objective risk assessment. In: Rainford D, Gradwell D, editors. Ernsting's aviation medicine. 4th ed. London (UK): Hodder-Arnold; 2006:547–566.
4. Federal Aviation Administration. 2021 Guide for Aviation Medical Examiners. [Accessed 20 June 2023]. 2021. Available from https://www.faa.gov/sites/faa.gov/files/about/office_org/headquarters_offices/avs/guide.pdf.
5. Gray G, Bron D, Davenport ED, d'Arcy J, Guettler N, et al. Assessing aeromedical risk: a three-dimensional risk matrix approach. Heart. 2019; 105(Suppl. 1):s9–s16.
6. Gray GW, Sargsyan AE, Davis JR. Clinical risk management approach for long-duration space missions. Aviat Space Environ Med. 2010; 81(12): 1128–1132.
7. International Civil Aviation Organization. Manual of civil aviation medicine. 2012. [Accessed 21 Dec. 2021]. Available from https://www.icao.int/publications/documents/8984_cons_en.pdf.
8. Janic M. An assessment of risk and safety in civil aviation. J Air Transp Manage. 2000; 6(1):43–50.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33(1):159–174.
10. Mitchell SJ, Evans AD. Flight safety and medical incapacitation risk of airline pilots. Aviat Space Environ Med. 2004; 75:260–268.
11. Mulloy A, Wielgosz A. Cardiovascular risk assessment in pilots. Aerosp Med Hum Perform. 2019; 90(8):730–734.
12. Olsson U. Maximum likelihood estimation of the polychoric correlation coefficient. Psychometrika. 1979; 44(4):443–460.
13. Prudhomme MB, Ropp LG, Sauer SW, LaVan JT. Aeromedical risk assessment of pharmaceuticals using evidence-based medicine. Aerosp Med Hum Perform. 2015; 86(9):824–829.
14. Schweder T, Hjort NL. Confidence, likelihood, probability, vol. 41. Cambridge (UK): Cambridge University Press; 2016:1–7.
15. Tunstall-Pedoe H. Acceptable cardiovascular risk in aircrew. Introduction. Eur Heart J. 1988; 9(Suppl. G):9–11.
16. Tunstall-Pedoe H. Risk of a coronary heart attack in the normal population and how it might be modified in flyers. Eur Heart J. 1984; 5(Suppl. A): 43–49.
17. U.S. Air Force. Risk assessment and acceptance. Washington (DC): Department of the Air Force; 2017. Air Force Airworthiness Bulletin 150B. [Accessed 12 April 2022]. Available from https://daytonaero.com/wp-content/uploads/AWB-150A.pdf.
18. U.S. Air Force. Risk management. Washington (DC): Department of the Air Force; 2019. Air Force Instruction 90-802. [Accessed 12 April 2022]. Available from https://static.e-publishing.af.mil/production/1/af_se/publication/afi90-802/afi90-802.pdf.
19. U.S. Air Force. Risk management guidelines and tools. Washington (DC): Department of the Air Force; 2013. Air Force Pamphlet 90-803 [later superseded by 2022 version].
20. U.S. Air Force. Risk management guidelines and tools. Washington (DC): Department of the Air Force; 2022. Air Force Pamphlet 90-803. [Accessed 12 April 2022]. Available from https://static.e-publishing.af.mil/production/1/af_se/publication/dafpam90-803/dafpam90-803.pdf.
21. U.S. Air Force. USAF airworthiness. Washington (DC): Department of the Air Force; 2022. Air Force Instruction 62-601. [Accessed 28 Oct. 2022]. Available from https://static.e-publishing.af.mil/production/1/saf_aq/publication/dafi62-601/dafi62-601.pdf.
22. U.S. Air Force. USAF airworthiness certification. Washington (DC): Department of the Air Force; 2019. Air Force Policy Document 62-6. [Accessed 12 April 2022]. Available from https://static.e-publishing.af.mil/production/1/saf_aq/publication/afpd62-6/afpd62-6.pdf.