Aeromedical Risk Assessment of Pharmaceuticals Using Evidence-Based Medicine

Michael B. Prudhomme; Lincoln G. Ropp; Samual W. Sauer; Joseph T. LaVan

BACKGROUND: Using concepts from evidence-based medicine, systems theory, and risk assessment, a standardized model was developed to accept or reject medications for use in flight. The model calculates the risk scores of medications, which can then be compared to an organization's acceptable risk tolerance.

METHODS: Risk scores for each medication were established by summing the products of incidence rates and severity scores for all published side effects. The incidence of each side effect was obtained in an evidence-based manner and each assigned a severity multiplier. Using statistical analysis of the calculated risk scores of approved medications, an acceptance control chart was generated.

RESULTS: Range of calculated risk scores of historically approved medications was 10–9140. Six Sigma Acceptance Control Line was calculated at 1.5 SDs above the mean and was 9822. Risk score range of medications generally felt unsafe was 27,010–41,294. Risk score range of medications under consideration for approval was 986–6863.

DISCUSSION: This novel approach to medication approval is the first in aerospace medicine to attempt to combine evidence-based medicine, risk analysis, and control charts to standardize and streamline the medication approval process within an organization. The model was validated by testing against medications generally accepted to be unsafe for use in flight. These medications fell several deviations above the control line. Other medications not yet authorized fall well below the acceptance line and could be considered for approval.

KEYWORDS: medication approval, evidence-based medicine, risk analysis, control charts, systems theory.

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.

A fundamental responsibility of aerospace medicine is the analysis and mitigation of the human component's risk to the aviation system. Part of this risk mitigation uses medications, when medically indicated, to treat or prevent disease in aviation personnel. However, medications potentially increase risk by inducing undesirable aeromedical effects, which have been shown to result in accidents. For instance, during a 16-yr period from 1990–2005, the Federal Aviation Administration detected antihistamines alone in 103 fatalities and in combination with other drugs and/or ethanol in an additional 235. Antihistamines were found in approximately 4% and 11% of the fatalities or accidents in 1990 and in 2004, respectively. The use of antihistamines was determined by the National Transportation Safety Board to be the cause of 13 and a factor in 50 of 338 accidents (18%).[2] Clearly, not all medications are safe for use during flight; the challenge is to determine which ones are.

Aviation organizations use varying processes to decide which medications are approved for flight. The Federal Aviation Administration uses a standing therapeutics committee to evaluate medications via an expert panel format. Similarly, each of the United States Armed Forces uses expert panels to consider medications and update their organizations' aeromedical policies. Though the panel may review Level I evidence, if done in a nonstandardized manner, each decision or a cumulative series of decisions may result in a Level V conclusion or grade D recommendation.[9] Furthermore, the constitution of the expert

From the Naval Aerospace Medicine Institute, Pensacola, FL.

This manuscript was received for review in January 2015. It was accepted for publication in May 2015.

Address correspondence to: LCDR Michael Prudhomme, M.D., Naval Aerospace Medical Institute, Naval Medicine Operational Training Center, Pensacola, FL32508; michael.b.prudhomme.mil@mail.mil.

DOI: 10.3357/AMHP.4273.2015

panel may fluctuate drastically over time and varying personalities may also play a powerful role in expert opinion. Group think is also another common pitfall that may occur in expert panels, potentially causing acquiescence in order to render a decision.[5] More notably, none of the organizations discussed above have published a standard method or the technical considerations with which they evaluate medications. Inherently inefficient as designed, these committees are currently charged with deciding the aeromedical acceptability for each individual medication. This may explain why a medication approved for use by one military service may be considered disqualifying in another service even though both operate similar aircraft in terms of capabilities and physiological conditions. Based on the historical record of published recommendations, these expert panels do not efficiently or rapidly analyze new medications. The approval processes, from initiation to decision, may take several years. Consequently, many medications with low aeromedical risk go unevaluated and remain prohibited from use due to lack of resources. This may prevent the use of appropriate medications which are indicated by the standard of care to improve or prevent disease and this is ethically problematic. These processes are cumbersome, time consuming, duplicative, and promulgate the publication of inconsistent conclusions across the aviation enterprise.

To improve current practice, it would be useful to develop a more evidence-based approach to aeromedical risk mitigation.[12] Ideally the evaluation process would be standardized, reproducible, evidence-based, efficient, responsive, and defensible, resulting in consistent evidenced-based aeromedical decisions. This paper presents just such a model. In the process of developing this model, we reviewed and incorporated best business practices from several disciplines. This included using concepts from evidence-based medicine, risk analysis, and systems theory.

Nearly every aeromedical policy decision includes use of risk analysis. Assessment of risk can be accomplished in many ways, but is usually divided into three primary subtasks:

- Risk Determination: involves identifying the risk and quantifying the probability of occurrence, and the severity of outcomes;
- Risk Measurement: analyzing data to quantify historical risk;
- Risk Evaluation: accepting or rejecting a determined or acceptable level of risk.[4]

Risk is broadly accepted as the product of the probability and the severity of an event. The probability of medication effects considered aeromedically unacceptable may be found by reviewing available Level I evidence in a systematic manner and applying it to each medication in a standard way. Randomized controlled trials (RCTs) describing the incidences (probability) of the adverse reactions of medications are widely available in published sources, especially for medications approved by the Federal Drug Administration in the past decade. However, determining the aeromedical risks of a medication is more challenging, as few trials include aviators as research subjects. There are even fewer published studies of medications actually evaluated in an RCT in the real or simulated physiological environment of flight. So the best available practice would be the use of nonaviator medication effects probability data. If and when aviator-specific data becomes available, this data can be substituted into the model.

Currently, there are no published thresholds of aeromedically acceptable medication risk on which to base risk acceptance or rejection. Adapting acceptable risk thresholds from statutory levels of De Minimis risk (for the FAA, unacceptable risk $= 10^{-6}$ failures per flight hour per component, broadly acceptable risk $= 10^{-9}$)[11] through the use of the 1% or 2% rule[8] is not practical or as potentially appropriate for medications as it is for medical diseases. First, the rule is based on knowing the annual incidence rate of an event such as the recurrence of a myocardial infarction after revascularization. In contrast, drug trials record each adverse reaction that occurred during the period of the study. These periods are frequently undisclosed in the open literature. Therefore, the incidence is not a true annual rate in terms of the proper epidemiological definition. Furthermore, the 1% rule makes several assumptions that are based on a second pilot taking control of the aircraft if the primary pilot becomes incapacitated during a critical phase of flight. However, in many military, helicopter, and general aviation flights the critical phases of flight may encompass the entire duration of the flight and not be limited to takeoff, departure, approach, and landing; in addition, many military aircraft are single pilot, so there is no one available to recover the flight if the primary pilot becomes incapacitated. Consequently, another approach is required to establish an acceptable risk threshold.

An organization's inherent risk threshold is associated with the perceived severity of the medication effects. The severity of an adverse reaction may be perceived differently by different organizations based on their unique operations, thus explaining why some medications are approved by one organization but not another. Therefore, a second option for establishing an acceptable risk threshold is to use precedent. Organizations have been making risk assessments of medications since the early 1900s. Though formal calculations are rarely reported, each organization apparently establishes an "acceptable" risk tolerance for medications by precedent. If this level of tolerance could be mathematically quantified, future policy decisions on medications could be made under the premise that new medications approved by the organization should not introduce more risk into the system than historically accepted.

Developing a mathematical model to calculate acceptable risk is a formidable undertaking and drawing on established concepts from other industries simplifies the process. Systems theory is one such concept which has already been applied to accident mitigation.[7] According to the theory, a system is "a set of interacting units or elements that form an integrated whole, intended to perform some function, or any structure that exhibits order, pattern and purpose, [which] implies some constancy over time. A system's purpose is the reason for its existence and

the starting point for measuring its success."[13] Envisioning aviation as a complex system allows application of this theory. In this case, the purpose of the system is to provide air transportation of people or payloads, and the goal is to provide this as safely and reliably as possible. In this system, medication use by aviators is one of the many interacting units that can affect how safely the system operates.

Control theory is a branch of systems theory. In control theory, a system is ether "in" or "out" of control. A system "in control" produces statistically consistent output, while a system "not in control" produces output that is less predictable. Statistically consistent is defined as output that falls within a certain number of standard deviations of the mean. Therefore, a system can produce an undesirable output, but still be "in control," i.e., mean not desirable, but all output within 1 SD. If a system is in control, then managers (in our case aeromedical policy makers) can manipulate the system input, such as medical and medication waivers, in order to achieve a desired output (an acceptable mishap rate). If a system is not in control, then input may produce unpredictable responses to the output, resulting in a system unable to be directly manipulated by managers.

Process control charts can be developed to visualize dynamic systems. These charts display statistical analysis of the output with reference to the process mean, to illustrate if the process is "in control." A control chart displays the process output on an axis with reference lines for the process mean and the upper (UCL) and lower control limits. These control boundaries are determined using statistics from historical output data.[1] Often these control limits are based on the Six Sigma (6σ) concepts of quality assurance, which state that a process is in control if the output constantly remains within ±3 SDs of the mean (±3 σ).[6] Another method is the use of an acceptance chart, which seeks to control a system within half the variance that can be derived from control charts by adding upper (UAL) and lower allowable limits. Allowable boundaries are often ±1.5 SD about the process mean, thus adding an increased element of quality assurance to process control.[3]

METHODS

PubMed and Google Scholar searches were conducted for published models concerning the approval of pharmaceuticals for use in flight. The phrases 'standardized aviation decision making; standardized aeromedical decision making for medications; aviation or flight medication approval; aeromedical risk assessment of pharmaceuticals or medication or drug' were used. Results were returned for analyzing individual medicines for use in aviation and models for the decision making process for individual medical conditions. However, no articles were found standardizing the process for analyzing multiple medications for use in aviators.

Using the constructs discussed in the introduction, we developed a mathematical model using published evidence and a standardized technique to calculate an objective risk score of any medication. The mathematical model was applied to currently approved medications, generating specific risk scores for each one and allowing for the calculation of the currently acceptable risk (average of these scores). The methodology was as follows:

1. Determine reference sample to establish historic risk threshold.
2. Identify all published adverse reactions of each medication.
3. Assign severity multiplier to each reaction based on recognized aeromedical concerns.
4. Develop a standardized protocol to establish the probability of adverse reactions.
5. Calculate aeromedical risk scores for reference medications.
6. Generate acceptance control chart.
7. Utilize chart to assess medications of interest to aeromedical concerns.

Thus we can evaluate a medication's acceptance or rejection for use in aviation based on its risk score in relation to the historically acceptable risk scores in naval aviation.

We applied the aforementioned methodology to U.S. Navy aeromedical practices in the following manner. The first step of this project was to determine a reference sample of historically approved medications on which to base the acceptable risk threshold. In an effort to minimize selection bias, data was extracted from a Department of Defense database which provides an electronic record of all prescriptions filled at Department of Defense facilities. The database was searched to extract all new prescriptions written for personnel on active duty in a flying status between 1 September 2011 and 1 September 2013.

From this data we identified the 70 medications most commonly prescribed to personnel in flight status, excluding those drugs not approved for use in flight by current Navy policy. **Table I** shows a sample of this data. The epidemiological analysis of the full data is currently ongoing for a separate publication.[10] In addition, we reviewed the waiver guide to ensure inclusion of at least one representative medication from each approved class of drugs. In this manner a

Table I. Top 15 Approved Medications Prescribed to U.S. Naval Aviators.

DRUG	NUMBER OF NEW PRESCRIPTIONS
Ibuprofen	1030
Loratadine	369
Acetaminophen	369
Fluticasone	355
Naproxen	291
Fexofenadine	259
Doxycycline	181
Augmentin	129
Valacyclovir	114
Esomeprazole	98
Amoxicillin	87
Meloxicam	73
Mometasone	69
Simvastatin	69
Ranitidine	63

reference sample of 31 medications by which to calculate acceptable risk was established.

Next, the frequencies of the side effects of each medication were obtained using a standardized and evidence-based approach. Open source and commonly available sources were used in the following order:

1. Lexicomp;
2. Daily Med;
3. Physician Desk Reference;
4. Drug trials submitted to the Federal Drug Administration; and
5. Other peer reviewed sources.

This method appears to provide the most accessible, reliable, and comprehensive Level 1 evidence. However, the prevalence of side effects is reported differently among the different sources. To further compound the issue, the prevalence is often reported as a range because the resources often combine data from multiple RCTs. In order to standardize the process for each medication the following procedure was utilized:

- Ranges were assigned a value equal to the arithmetic mean of the range.
- When data was reported as less than ($<$ x%), but a study of lower precedence provided either a percentage or a range, the lower precedence data was used for the side effect prevalence.
- When data in all studies was reported as less than a low percentage (often $<$ 1% or $<$ 2%), the side effect was scored as zero, as no mean could be calculated.
- When studies indicated a side effect was less than with placebo it was scored as zero.

Using generally accepted concepts of what constitutes an aeromedically adverse medication effect, each adverse reaction was placed in a severity category, and each category was assigned a severity score multiplier (**Table II**).

The next step was to calculate the risk score for each medication in the reference sample from the prevalence and severity data obtained from above. The risk of an event is generally considered to be the product of the likelihood and the severity of the event. This concept was then extrapolated to the risk of medication use. In order to account for multiple side effects of a single medication, we developed a formula to quantify the composite risk for a medication. This composite risk calculation is the sum of the prevalence of every published side effect multiplied by its perceived severity, and is represented algebraically by the formula below:

$$\sum\left(P\,\text{side effect}_a \times \text{severity side effect}_a\right) + \left(P\,\text{side effect}_b \times \text{severity side effect}_b\right)\ldots_{©}$$

We then calculated the average risk score for the reference sample. This was determined to be 3379, a unit-less number. This number represents a quantification of the average risk in medication use historically accepted by the organization under consideration. Though the boxplot of these scores appears to have a slight positive skew (**Fig. 1**), the Shapiro-Wilk normality test confirmed a normal distribution (0.929 $>$ 0.05) on which to base the other calculations. The UAL (+1.5 SDs = 8086) and UCL (+3 SDs = 12,739) were calculated from the mean and standard deviation of the reference sample (**Fig. 2**). In order to test the validity of the model, risk scores from medications known to be causal in mishaps were calculated, as were the scores of medications felt to be safe by other aeromedical organizations, but not approved by the Navy (Fig. 2).

Table II. Aeromedically Adverse Medication Effects.

AEROMEDICALLY ADVERSE MEDICATION EFFECTS	EXAMPLE	SEVERITY SCORE MULTIPLIER
Totally Incapacitating	Seizure	1000
Subtly Incapacitating	Drowsiness	100
Distracting	GERD	10
Mildly Distracting	Dry Mouth	1
No Aeromedical Consequence	Elevated LFT	0

GERD: gastroesophageal reflux disease; LFT: liver function test.

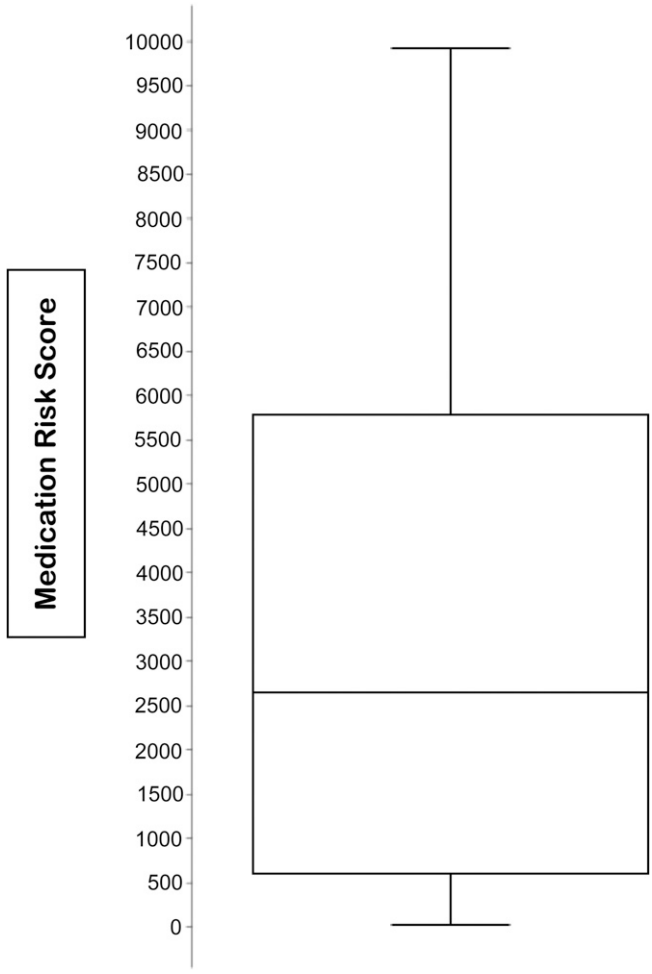

Fig. 1. Box plot of risk scores of historically approved medications.

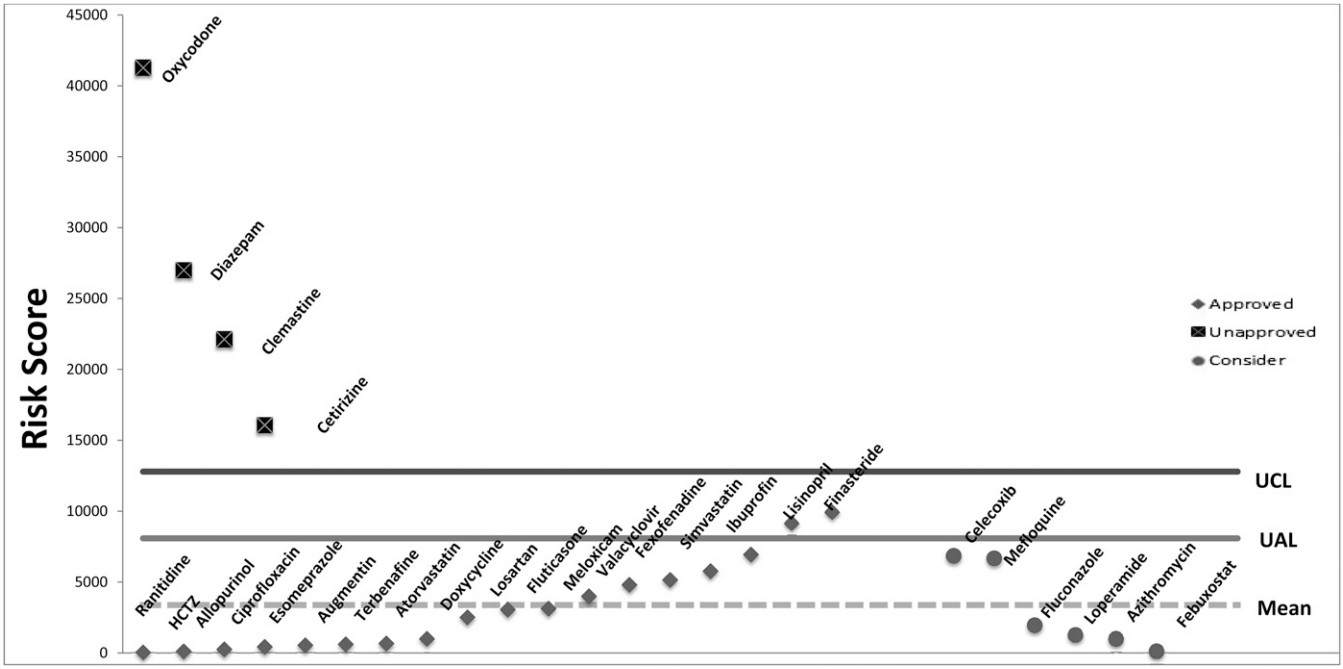

Fig. 2. Acceptance chart displaying approved, unapproved, and potentially acceptable medications.

RESULTS

Medications felt to be aeromedically unacceptable included a narcotic (oxycodone, 41,294), a benzodiazepine (diazepam, 27,010), a sedating H1 blocker (clemastine, 22,137), and a "minimally" sedating H1 blocker which is considered by many organizations to be too sedating for use in aviation (cetirizine, 16,084). All of these medications fell well above the UCL of 12,739. Several medications not approved by the U.S. Navy, but approved by other organizations fell below the UAL of 8086: mefloquine (6686), fluconazole (1951), azithromycin (986), loperamide (1255), and celecoxib (6863) (Fig. 2).

DISCUSSION

As developed, this model provides a rapid and evidence based method for evaluating the relative safety of medications for use in the aerospace environment. As a proof of concept, for medications with prevalence data readily available in one of the online databases, we were able to calculate risk scores of the medication in approximately 10 min. This allows one individual, using evidence based sources, to analyze and recommend a determination for many different medications in a much shorter time frame than the more cumbersome and nonstandardized methods currently used throughout aerospace medicine.

This model appears to be valid within current U.S. Navy aeromedical standards. Calculated risk scores appear consistent with the aeromedical status of various medications. Specifically, risk scores for medications found causal in published historic mishaps greatly exceeded the UCL. This is consistent with the application of the UCL to the threshold of aeromedically unacceptable risk. Most medications currently approved under U.S. Navy aeromedical standards fell below or near the UAL, also validating this point. The value of this tool in aeromedical decision making is pointed out by the fact that several medications approved by other aeromedical organizations (but not yet approved by the U.S. Navy) did not exceed the UAL. Under this model, a risk score below the UAL would be thought of as broadly acceptable risk and could identify a medication that should be considered for approval. Finally, medications under consideration with a calculated risk score falling between the UAL and the UCL may have tolerable risk, but careful consideration of the risk/benefit ratio and the possibility of alternative treatments must be given.

An inherent weakness of the model is that the prevalence data for adverse reactions is extrapolated from studies on the general population, but not yet determined in aviators and the aviation environment. However, when evaluating this model, it is important to recognize that this is the same level of evidence available and used when making the previous aeromedical decisions. The applicability of the data is no worse in the model; it is just being applied in a more objective, quantifiable fashion. In the future if higher levels of evidence were obtained in targeted studies in the aviation environment, substituting this specific data and recalculating the risk score is simple. This weakness is also a strength, as the best available prevalence data is from the general population and is thus generalizable to multiple organizations and industries.

Another limitation is that the severity scores for each medication, while standardized, are still subjective. From a practical standpoint, the perception of the severity of particular effects may be different among different industries or even among

different aeromedical organizations. Organizations with single-seat high-performance aircraft may have a very different perception of risk than those with multipiloted heavy or rotary wing aircraft. It is paramount to note when using this model, each organization must make a determination of the relative severity of each adverse reaction and assign an appropriate multiplier for use throughout their methodology. Furthermore, the model does not negate proper aeromedical judgment. As it relies primarily on phase III drug trials, it would be prudent to provide a sufficient observation period in the general population before considering new drugs for approval. While great effort was made to create a representative sample, the model could be further validated with larger data sets in the future.

Strengths of the model include simplicity, standardization, and efficiency. This method, using widely available resources, provides a simple technique for determining risk scores. It is standardized, evidence-based, defensible, and remains consistent with historic decisions. The protocol allows efficient evaluation of medications, enabling an organization to stay current with clinical practice guidelines in a highly dynamic profession. By calculating risk scores of all medications within the same class, it is possible to determine which one would carry the lowest aeromedical risk and validate the decision in a more objective fashion. Finally, the model is both reproducible and adaptable to other organizations and industries. Any other organization in any other industry may adapt the model by adjusting the severity scores for their specific environment and then following the procedures outlined above.

ACKNOWLEDGMENTS

The views expressed in this article are those of the author(s) and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the United States Government.

Authors and affiliation: Michael B. Prudhomme, M.D., M.P.H., Lincoln G. Ropp, B.S., D.O., and Joseph T. LaVan, M.D., M.P.H., Naval Aerospace Medical Institute, Pensacola, FL, and Samual W. Sauer, M.D., M.P.H., U.S. Army School of Aviation Medicine, Ft. Rucker, AL.

REFERENCES

1. American Society for Quality. Control chart. [Accessed Oct. 13, 2014]. Available from: http://asq.org/learn-about-quality/data-collection-analysis-tools/overview/control-chart.html.
2. Sen A, Akin A, Craft KJ, Canfield DV, Chaturvedi AK. First-generation H1 antihistamines found in pilot fatalities of civil aviation accidents, 1990-2005. Aviat Space Environ Med. 2007; 78(5):514–522.
3. Holmes D, Mergen A, editors. Building an acceptance chart around the process mean. Northeast Decision Sciences Institute Proceedings; March 26-28, 2010. Alexandria (VA): RIT Scholar Works; 2010.
4. Janic M. An assessment of risk and safety in civil aviation. Journal of Air Transport Management. 2000; 6(1):43–50.
5. Kerkering JC. Eliciting and analyzing expert judgment, a practical guide. Technometrics. 2002; 44(2):193–194.
6. Keller P. Six sigma demystified, 2nd ed. New York (NY): McGraw-Hill; 2011:3–19.
7. Leveson N. Applying systems thinking to analyze and learn from events. Saf Sci. 2011; 49(1):56–64.
8. Mitchell SJ, Evans AD. Flight safety and medical incapacitation risk of airline pilots. Aviat Space Environ Med. 2004; 75(3):260–268.
9. Oxford Centre for Evidence-Based Medicine. Levels of evidence (March 2009). [Accessed April 22, 2015]. Available from: http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009.
10. Prudhomme M. A model for aeromedical risk assessment of pharmaceuticals, utilizing evidence based medicine and acceptance control charts [Thesis]. Pensacola (FL): University of West Florida; 2013.
11. Scarlett LL, Linkov L, Kousky C. Risk management practices: cross-agency comparisons with minerals management service. Washington (DC): RFF Press; 2011.
12. Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. BMJ. 1999; 318(7183):593–596.
13. Skyttner L. General systems theory: ideas and applications. Singapore: World Scientific; 2001.