Modeling Predictors of Duties Not Including Flying Status

Anthony P. Tvaryanas; Converse Griffith Jr.

INTRODUCTION: The purpose of this study was to reuse available datasets to conduct an analysis of potential predictors of U.S. Air Force aircrew nonavailability in terms of being in "duties not to include flying" (DNIF) status.

METHODS: This study was a retrospective cohort analysis of U.S. Air Force aircrew on active duty during the period from 2003–2012. Predictor variables included age, Air Force Specialty Code (AFSC), clinic location, diagnosis, gender, pay grade, and service component. The response variable was DNIF duration. Nonparametric methods were used for the exploratory analysis and parametric methods were used for model building and statistical inference.

RESULTS: Out of a set of 783 potential predictor variables, 339 variables were identified from the nonparametric exploratory analysis for inclusion in the parametric analysis. Of these, 54 variables had significant associations with DNIF duration in the final model fitted to the validation data set. The predicted results of this model for DNIF duration had a correlation of 0.45 with the actual number of DNIF days. Predictor variables included age, 6 AFSCs, 7 clinic locations, and 40 primary diagnosis categories.

DISCUSSION: Specific demographic (i.e., age), occupational (i.e., AFSC), and health (i.e., clinic location and primary diagnosis category) DNIF drivers were identified. Subsequent research should focus on the application of primary, secondary, and tertiary prevention measures to ameliorate the potential impact of these DNIF drivers where possible.

KEYWORDS: aircrew, availability, duties not to include flying, DNIF, epidemiology.

Tvaryanas AP, Griffith C Jr. *Modeling predictors of duties not including flying status.* Aerosp Med Hum Perform. 2018; 89(1):52–57.

Practitioners of clinical medicine are trained to prevent, diagnose, and treat conditions that alter a patient's physiology and functional state in a normal environment. Practitioners of aerospace medicine must also understand the interaction of a patient's normal or abnormal physiology and functional state within the mission environment and the resulting impact on overall flight safety and performance. Accordingly, in managing acute and chronic illnesses, the aerospace medicine practitioner has the additional duty of rendering an aeromedical disposition; that is, an occupational medicine determination whether a particular aircrew member is "fit to fly." Prudent aerospace medicine practitioners also track the epidemiology of conditions that limit aircrew availability and work toward prevention of these conditions.[5] Given ever-present resource constraints, not the least of which is aerospace medicine practitioner time, prevention efforts should focus on those conditions that are the primary driver of aircrew nonavailability. Unfortunately, there is scant published literature on this subject to inform the aerospace medicine practitioner.

The purpose of this study was to reuse available datasets to conduct an exploratory analysis of potential predictors of U.S. Air Force (USAF) aircrew nonavailability in terms of being in "duties not to include flying" (DNIF) status. The following hypotheses guided this study:

- H_1: Demographic factors, including age and gender, are associated with duration of DNIF status.
- H_2: Occupational factors, including Air Force Specialty Code (AFSC), service component, and pay grade, are associated with duration of DNIF status.
- H_3: Health factors, in terms of diagnoses and clinic, are associated with duration of DNIF status.

From the 711th Human Performance Wing, Wright-Patterson AFB, OH.

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

Address correspondence to: Anthony P. Tvaryanas, 9685 Olde Georgetown Way, Centerville, OH 45458; aptvaryanas@att.net.

DOI: https://doi.org/10.3357/AMHP.4980.2018

METHODS

Study Design

This study was conducted under a human-use protocol approved by the 711th Human Performance Wing Institutional Review Board. A waiver of informed consent of participants was granted due to the impracticality of obtaining written consent from each participant in the study population. This study was a retrospective cohort analysis of USAF aircrew on active duty during the period from 2003–2012. This study reused a dataset created for a study analyzing all outpatient healthcare encounters occurring in any of the USAF's Flight and Operational Medicine Clinics (FOMCs) during the period from 2003–2012[7] as well as archival data on DNIF events extracted from the Aeromedical Services Information Management System. Inclusion criteria were USAF service members receiving care at a FOMC with at least one DNIF episode. Participants were excluded if they had missing data in the response variable, DNIF, or in the personal identifier.

Data and Variables

The basic unit of analysis was a DNIF episode. The duration of the DNIF episode and the associated primary diagnosis, recorded in terms of International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes, were obtained from the Aeromedical Services Information Management System. Participant age (continuous), gender (categorical with 2 levels), pay grade (categorical with 16 levels), AFSC [categorical with 270 levels (using career group, career field, and career field subdivision for enlisted personnel and career group and functional area for officers)], service component (categorical with 3 levels), and FOMC location (categorical with 77 levels) for each DNIF episode were obtained from the pre-existing study dataset; details on the creation of this dataset are available elsewhere.[7]

Diagnosis codes were recoded using a software tool developed as part of the Healthcare Cost and Utilization Project. The Clinical Classification Software for ICD-9-CM aids analysts to collapse diagnostic data from over 14,000 diagnosis codes that make up the ICD-9-CM standardized coding system into clinically meaningful categories.[1] The 367 tertiary level classifications were used, with 22 additional levels for Department of Defense specific categories, such as "Medication Education," "Armed Forces Health Exam," and "Travel Medication Education," for a total of 389 levels.

Statistical Analysis

The original, reused dataset[7] comprised 90,331 distinct participants. A total of 7858 participants did not meet the study inclusion criteria or were excluded because of missing data. The final study population comprised 389,976 DNIF events from 82,473 distinct participants. The study dataset was randomly partitioned into several samples: a learning sample (235,919 DNIF events from 50,000 distinct participants) for exploratory analysis and initial variable selection, a training sample (70,150 DNIF events from 15,000 distinct participants) for further variable selection using a marginal longitudinal model, and a validation sample (71,938 DNIF events from 15,000 distinct participants) for statistical inference using the variables selected in the first two steps. A fourth remainder sample of 11,969 DNIF events from 2473 distinct participants was unused; this sample was held back in case further data exploration was necessary. Nonparametric methods were used for the exploratory analysis and parametric methods were used for model building and statistical inference given the greater ease of interpretation of the latter (e.g., standard errors, *P*-values, etc.). Separating variable selection and model building ensured that the reported standard errors and *P*-values were valid.

Tree-based gradient boosting machine (GBM)[3] modeling was used for exploratory analysis on the learning sample. The GBM variable importance capability was used to select the most influential predictors; larger variable importance scores suggested greater importance in terms of predicting the response. Prior to analysis, all high-level categorical variables were one-hot encoded; that is, a separate dummy variable was created for every level of each variable. This procedure yielded a total of 783 predictor variables that were used for exploratory analysis. Variables with nonzero importance scores were subsequently included in the parametric analyses. Since the study objective was to identify population-wide predictors of DNIF duration rather than inference on individuals, a marginal model rather than a longitudinal model was used.[2] A negative binomial model with a log link function was chosen because the response variable was a count variable with dissimilar mean and variance (thus making a Poisson model a suboptimal choice). Predictor variables included age, gender, pay grade, AFSC, service component, FOMC location, and diagnosis. Participant was a random repeated measure in the model and a compound symmetry (exchangeable) covariance structure was assumed.

R version 3.3.2[6] was used for data preparation and calculation of summary statistics. The R gbm package, version 2.1.1, was used to accomplish the GBM modeling. SAS version 9.4 (SAS Institute, Cary, NC) was used to create the sample datasets (Proc SurveySelect). SAS (Proc GenMod) was used to fit the marginal longitudinal model on the training sample and estimate the model of the validation sample. Statistical significance was defined as $P = 0.0001$.

RESULTS

Table I provides descriptive statistics for the measured variables for the final study population. Clinic location is not displayed, as only clinic pseudoidentifiers were provided to preserve data de-identification. With the exception of primary diagnosis category, summary statistics were computed on the basis of the population of unique participants using a randomly selected DNIF event during the first year of observation to establish a measurement for each variable. In contrast, summary statistics for primary diagnosis category were computed based on the population of DNIF events.

In the final validation sample, the predicted results of the marginal longitudinal model for DNIF duration had a correlation (*r*)

Table I. Descriptive Statistics for the Study Population.

VARIABLE	DESCRIPTIVE STATISTIC
N	82,473
DNIF events:	
N	389,976
Duration, days, median (IQR)	7.00 (13.58)
Age, yr, median (IQR)	27 (12)
Gender (ref male), no. (%)	72, 834 (88.31)
Service component, no. (%):	
Active duty	75,379 (91.40)
Reserve	3909 (4.74)
National Guard	3185 (3.86)
AFSC, no. (%):	
Officer:	
92TX Pilot trainee	7333 (20.59)
11FX Fighter pilot	3351 (9.41)
13SX Space and missile	2824 (7.93)
11MX Mobility pilot	2697 (7.57)
11AX Airlift pilot	2068 (5.82)
11KX Trainer pilot	1747 (4.91)
13BX Air battle manager	1570 (4.41)
11TX Tanker pilot	947 (2.67)
11RX Reconnaissance/surveillance/ electronic warfare pilot	784 (2.20)
12RX Reconnaissance/surveillance/ electronic warfare combat systems officer	783 (2.20)
12BX Bomber combat systems operator	713 (2.00)
11BX Bomber pilot	629 (1.77)
11SX Special operations pilot	563 (1.58)
12FX Fighter combat systems officer	561 (1.57)
62EX Developmental engineer	557 (1.56)
46FX Flight nurse	486 (1.36)
11HX Helicopter pilot	481 (1.35)
92SX Student officer authorization	476 (1.34)
Other*	7040 (19.77)
Enlisted:	
1C1XX Air traffic control	4773 (13.06)
1A2XX Aircraft loadmaster	4233 (11.58)
9T0XX Basic enlisted airman	2764 (7.56)
1A1XX Flight engineer	2344 (6.41)
1A8XX Airborne cryptologic linguist	2303 (6.30)
1A3XX Airborne mission system	1718 (4.70)
1C6XX Space systems operations	1469 (4.02)
1A0XX In-flight refueling	1447 (3.96)
4N0XX Aerospace medical service	1313 (3.59)
1A4XX Airborne operations	1259 (3.44)
1C5XX Command and control battle management operations	1118 (3.06)
2A5XX Aerospace maintenance	818 (2.24)
3P0XX Security forces	802 (2.19)
1C4XX Tactical air control party	749 (2.05)
1C2XX Combat control	730 (2.00)
1T2XX Pararescue	729 (1.99)
1N1XX Geospatial intelligence	569 (1.56)
1A7XX Aerial gunner	513 (1.4)
Other*	6906 (18.91)
Missing	10,306 (12.50)
Primary diagnosis category,† no. (%):	
Diseases of the respiratory system	104,637 (26.83)
DoD specific: education or counseling	48,117 (12.34)
Diseases of the digestive system	31,177 (7.99)
Diseases of the nervous system and sense organs	30,625 (7.85)
Symptoms; signs, ill-defined conditions and factors influencing health status	26,360 (6.76)
Diseases of the musculoskeletal system and connective tissue	24,521 (6.29)
Injury and poisoning	22,404 (5.74)
Diseases of the skin and subcutaneous tissue	5529 (1.42)
Infectious and parasitic diseases	5425 (1.39)
Residual codes, unclassified, all E codes	5228 (1.34)
Complications of pregnancy; childbirth; and the puerperium	4917 (1.26)
Diseases of the genitourinary system	4899 (1.26)
Diseases of the circulatory system	4277 (1.10)
Mental illness	3494 (0.90)
Endocrine; nutritional; and metabolic diseases immunity disorders	3129 (0.80)
DoD specific exams	1228 (0.31)
Neoplasms	768 (0.20)
Congenital anomalies	462 (0.12)
Diseases of the blood and blood-forming organs	174 (0.04)
Certain conditions originating in the perinatal period	83 (0.02)
Other DoD specific diagnoses	16 (< 0.01)
DoD specific: traumatic brain injury	5 (< 0.01)
Missing	62,501 (16.03)
Pay grade, no. (%):	
Officer	
O1	16,433 (36.29)
O2	5180 (11.44)
O3	11,057 (24.42)
O4	6045 (13.35)
O5	5178 (11.44)
O6+	1388 (3.07)
Enlisted	
E1	1688 (4.54)
E2	2366 (6.36)
E3	10,809 (29.06)
E4	5152 (13.85)
E5	8319 (22.37)
E6	4935 (13.27)
E7	2979 (8.01)
E8	753 (2.02)
E9	191 (0.51)

IQR = interquartile range; DoD = Department of Defense.
* Only AFSCs comprising 80% of participants shown for brevity.
† Healthcare Cost and Utilization Project-Clinical Classification Software secondary level diagnosis categories shown for brevity.

of 0.45 with the actual number of DNIF days. Out of the 783 predictor variables used in the GBM model fitted on the learning data, 339 variables had a nonzero relative influence and were included in the parametric analyses. Of these predictor variables, 84 variables exclusive of the intercept had statistically significant associations at $P \leq 0.0001$ with DNIF duration when the initial negative binomial model was fitted using the training data. There were 53 variables, not including the intercept, which had statistically significant associations at alpha $P \leq 0.0001$ with DNIF duration when the final negative binomial model was fitted on the validation dataset (**Table II**).

Based on the model results, we partially accept hypothesis 1 that demographic factors are associated with the duration of DNIF status. There was a significant association with age and duration of DNIF status, while gender was not a predictor

Table II. Negative Binomial Regression Model Results for DNIF Duration.

VARIABLE	EXPECTED DAYS DNIF	*P*-VALUE
Intercept	15.50	<0.0001
Age	15.84	<0.0001
Gender (ref = male)	17.16	0.0003
Clinic location:		
CLID025606592	17.05	0.0925
CLID068902320	16.90	0.1810
CLID093047257	20.81	<0.0001
CLID106868943	20.07	0.0001
CLID109616999	18.97	0.0010
CLID110435376	28.34	<0.0001
CLID111016652	8.74	<0.0001
CLID125682959	24.51	<0.0001
CLID147851280	24.03	0.0066
CLID254322565	22.98	<0.0001
CLID322301264	17.66	0.0555
CLID381735261	37.92	<0.0001
Primary diagnosis category:		
Acute bronchitis	7.15	<0.0001
Administrative/social admission	10.90	0.0001
Allergic reactions	10.60	0.0002
Bipolar disorders	57.85	<0.0001
Blindness and vision defects	21.07	<0.0001
Calculus of urinary tract	27.57	<0.0001
Cardiac dysrhythmias	44.32	<0.0001
Cataract	46.23	<0.0001
Cellulitis and abscess	6.93	<0.0001
Codes related to mental health disorders	43.25	<0.0001
Coronary atherosclerosis and other heart disease	41.63	<0.0001
Depressive disorders	62.41	<0.0001
DoD specific: medication education	6.71	<0.0001
Ectopic pregnancy	80.02	<0.0001
Encephalitis, except that caused by TB or STD	6.35	0.0066
Endometriosis	26.53	0.1352
Essential hypertension	26.02	<0.0001
Fracture of lower limb	39.19	<0.0001
Fracture of upper limb	32.25	<0.0001
Gastritis and duodenitis	6.59	<0.0001
Glaucoma	26.37	0.0012
Heart valve disorders	32.76	0.0421
Influenza	6.03	<0.0001
Inguinal hernia	25.43	<0.0001
Intervertebral disc disorders	54.22	<0.0001
Migraine	50.80	<0.0001
Nausea and vomiting	4.72	<0.0001
Non-Hodgkins lymphoma	19.09	0.4444
Other abdominal hernia	24.68	<0.0001
Other aftercare	18.13	0.0002
Other and ill-defined heart disease	76.58	0.0061
Other and unspecified gastrointestinal disorders	7.62	<0.0001
Other and unspecified asthma	58.14	0.0026
Other chronic pulmonary disease	5.96	<0.0001
Other complications of pregnancy	60.39	<0.0001
Other fractures	33.31	<0.0001
Other mycoses	7.47	<0.0001
Other nontraumatic joint disorders	22.71	<0.0001
Other thyroid disorders	57.01	<0.0001
Other upper respiratory infections	5.68	<0.0001
Other viral infections	6.07	<0.0001
Otitis media and related conditions	8.35	<0.0001
Outcome of delivery (V codes)	85.02	<0.0001
Peri-, endo-, & myocarditis; cardiomyopathy (except that caused by TB or STD)	23.91	0.1869
Phlebitis; thrombophlebitis and thromboembolism	76.66	<0.0001
Pneumonia (except that caused by TB or STD)	9.43	0.0011
Pulmonary heart disease	56.11	0.0008
Regional enteritis & ulcerative colitis	56.90	<0.0001
Residual codes; unclassified; all E codes	20.16	<0.0001
Retinal detachments; defects; vascular occlusion; and retinopathy	37.83	<0.0001
Spondylosis and allied disorders	43.09	0.0060
Sterilization	7.60	<0.0001
Urinary tract infections	6.05	<0.0001
AFSC:		
11EX Experimental test pilot	12.65	0.4116
11FX Fighter pilot	11.12	<0.0001
11KX Trainer pilot	11.96	<0.0001
11MX Mobility pilot	12.25	<0.0001
11RX Reconnaissance/surveillance/electronic warfare pilot	11.56	<0.0001
1C1 × 1 Air traffic control	10.47	<0.0001
1C2 × 0 Combat control	24.66	0.2396
1C3 × 1 Command post	15.31	0.9798
1N3 × 4 Cryptologic language analyst	33.63	0.0313
21RX Logistics readiness	19.81	0.5358
3C0 × 1 Communication-computer systems	12.27	0.4961
3E5 × 1 Engineering	9.21	0.0686
44AX Chief, hospital/clinic services	29.75	0.0136
48AX Aerospace medicine specialist	13.03	0.2637
4E0 × 1 Public health	5.05	<0.0001
83RX Recruiting service	15.68	0.9685
91WX Wing commander	11.47	0.0590

CLID = clinic identifier; DoD = Department of Defense; STD = sexually transmitted disease; TB = tuberculosis.

variable selected for inclusion in the model. We also partially accept hypothesis 2 that occupational factors are associated with the duration of DNIF status. Of the occupational factors considered, only AFSC was selected for inclusion in the model, and then only 6 of the potential 270 levels of this variable were included and significant in the final model. We accept hypothesis 3 that health factors are associated with the duration of DNIF status. Of 389 diagnostic categories, 40 were included and significant in the final model, while 7 out of 77 potential clinic locations were included. **Fig. 1** provides a Pareto display of the primary diagnosis categories that were significantly associated with expected days DNIF. Based on observed effect size, 6 clinics and 25 diagnosis categories were the primary drivers of DNIF duration.

DISCUSSION

To the best of the authors' knowledge, this study is the first attempt to systematically explore potential predictors of USAF aircrew nonavailability in terms of being in DNIF status over a 10-yr period. Significant associations were observed between

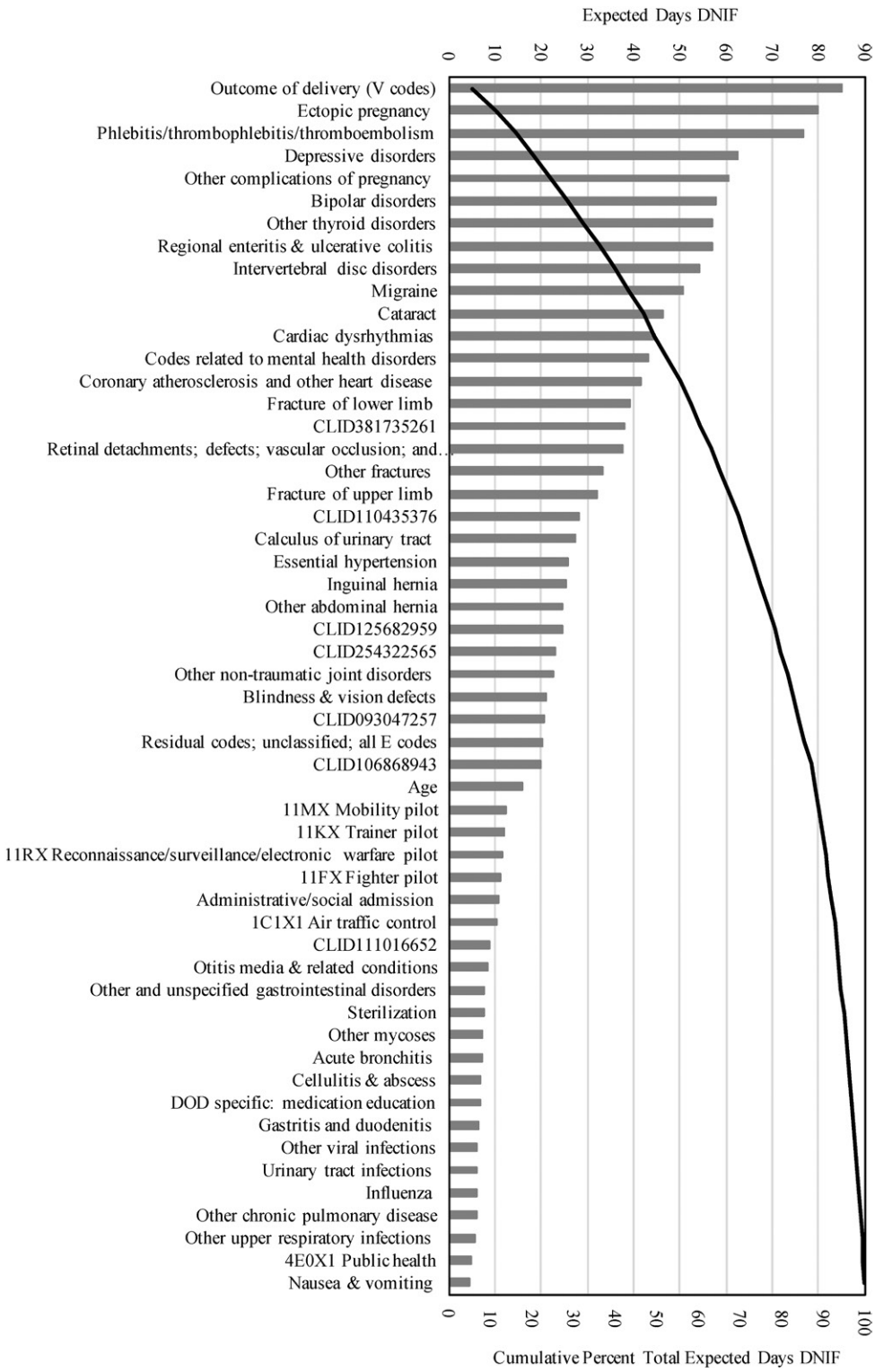

Fig. 1. Pareto display of the significant predictors of expected DNIF duration.

age, AFSC, clinic, and primary diagnosis category and expected DNIF duration. While controlling for specific diagnoses, increasing age was positively associated with expected DNIF duration. Six AFSCs were associated with an increased expected DNIF duration; however, these AFSCs were not significant drivers of DNIF duration based on the Pareto analysis. There was observed variability in expected DNIF duration based on clinic, with six clinics identified as significant DNIF drivers after controlling for other demographic, occupational, and health factors. As anticipated, multiple primary diagnosis categories were associated with increased expected DNIF duration. Based on Pareto analysis, 25 of these diagnosis categories appeared to be significant DNIF drivers relative to the other diagnoses: reproductive/pregnancy-related conditions, mental health conditions, fractures and degenerative joint conditions, cardiopulmonary conditions, ocular conditions, thyroid disorders, migraine headaches, enteritis and colitis, hernias, and renal calculi. Of note, gender was not associated with expected DNIF duration after controlling for diagnoses.

Given this analysis, the next step is to evaluate those conditions found to be significant DNIF drivers and identify opportunities for primary, secondary, and tertiary prevention.[4] Since infectious diseases were not among the DNIF drivers, traditional primary prevention measures focusing on vaccination are of limited utility. Instead, primary prevention should focus on those conditions caused by injuries and/or toxic exposures resulting from modifiable environmental exposures. Secondary prevention should focus on screening, either for specific conditions or antecedent, modifiable risk factors for those conditions (e.g., hypertension and coronary atherosclerotic disease). Routine screening is already accomplished as part of the mandated, annual periodic health assessment. Subsequent research, however, is needed to correlate current screening tools with observed DNIF drivers. Finally, tertiary prevention activities should seek to minimize expected DNIF duration after a condition occurs by optimizing treatment selection and delivery throughout the care cycle for the condition.

In terms of health informatics, the study methodology suggests a future approach for creating a near real time dashboard

or predictive software to support more active management of aircrew availability. This study was conducted by accessing already existing data collected for organizational business purposes and available from centralized databases. Accordingly, the aerospace medicine community of practice should explore analytic software solutions for connecting, analyzing, and visualizing this data to uncover patterns and predict trends, thereby better realizing the value inherent in the data in terms of optimizing aircrew availability.

In conclusion, specific demographic (i.e., age), occupational (i.e., AFSC), and health (i.e., clinic location and primary diagnosis category) factors were identified that were significantly associated with expected DNIF duration. Subsequent research should focus on the application of primary, secondary, and tertiary prevention measures to ameliorate the potential impact of these DNIF drivers where possible.

ACKNOWLEDGMENTS

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Air Force, the Department of Defense, or the U.S. Government.

Authors and affiliations: Anthony P. Tvaryanas, M.D., Ph.D., KBRwyle, and Converse Griffith, Jr., M.S., Ph.D., JYG Innovations, Dayton, OH.

REFERENCES

1. Agency for Healthcare Research and Quality. HCUP tools and software. Healthcare Cost and Utilization Project (HCUP). 2017. [Accessed 27 May 2017]. Available from http://www.hcup-us.ahrq.gov/tools_software.jsp.
2. Fitzmaurice GM, Laird NM, Ware JH. Applied longitudinal analysis, 2nd ed. Hoboken (NJ): John Wiley & Sons, Inc.; 2011.
3. Friedman JH. Stochastic gradient boosting. Comput Stat Data Anal. 2002; 38(4):367–378.
4. Jekel JF, Katz DL, Elmore JG, Wild DMG. Epidemiology, biostatistics, and preventive medicine, 3rd ed. Philadelphia (PA): Saunders/Elsevier; 2007.
5. Orford RR, Silberman WS. Pilot health and aeromedical certification. In: Davis JR, Johnson R, Stepanek J, Fogarty JA, editors. Fundamentals of aerospace medicine, 4th ed. Philadelphia (PA): Lippincott Williams and Wilkins; 2008:279–305.
6. R Core Team. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2013. Available from https://www.R-project.org/.
7. Tvaryanas AP, Maupin GM, Fouts BL. A 10-year cross-sectional analysis of Air Force Flight and Operational Medicine Clinic health care services. Mil Med. 2016; 181(5):424–433.