Accident Analysis and Prevention 57 (2013) 49–54

Contents lists available at SciVerse ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

US commercial air tour crashes, 2000–2011: Burden, fatal risk factors, and FIA Score validation

Sarah-Blythe Ballard[a,b,*], Leland P. Beaty[c], Susan P. Baker[c]

[a] *Johns Hopkins Bloomberg School of Public Health, Department of International Health, United States*
[b] *United States Navy, United States*
[c] *Johns Hopkins Center for Injury Research and Policy, United States*

ARTICLE INFO

Article history:
Received 29 October 2012
Received in revised form 19 March 2013
Accepted 21 March 2013

Keywords:
Commercial air tours
Air tours
Sightseeing
FIA
FIA Score

ABSTRACT

Introduction: This study provides new public health data concerning the US commercial air tour industry. Risk factors for fatality in air tour crashes were analyzed to determine the value of the FIA Score in predicting fatal outcomes.
Methods: Using the Federal Aviation Administration's (FAA) General Aviation and Air Taxi Survey and National Transportation Safety Board data, the incidence of commercial air tour crashes from 2000 through 2010 was calculated. Fatality risk factors for crashes occurring from 2000 through 2011 were analyzed using regression methods. The FIA Score, Li and Baker's fatality risk index, was validated using receiver operating characteristic (ROC) curves.
Results: The industry-wide commercial air tour crash rate was 2.7 per 100,000 flight hours. The incidence rates of Part 91 and 135 commercial air tour crashes were 3.4 and 2.3 per 100,000 flight hours, respectively (relative risk [RR] 1.5, 95% confidence interval [CI] 1.1–2.1, $P = 0.015$). Of the 152 air tour crashes that occurred from 2000 through 2011, 30 (20%) involved at least one fatality and, on average, 3.5 people died per fatal crash. Fatalities were associated with three major risk factors: fire (adjusted odds ratio [AOR] 5.1, 95% CI 1.5–16.7, $P = 0.008$), instrument meteorological conditions (AOR 5.4, 95% CI 1.1–26.4, $P = 0.038$), and off-airport location (AOR 7.2, 95% CI 1.6–33.2, $P = 0.011$). The area under the FIA Score's ROC curve was 0.79 (95% CI 0.71–0.88).
Discussion: Commercial air tour crash rates were high relative to similar commercial aviation operations. Disparities between Part 91 and 135 air tour crash rates reflect regulatory disparities that require FAA action. The FIA Score appeared to be a valid measurement of fatal risk in air tour crashes. The FIA should prioritize interventions that address the three major risk factors identified by this study.

Published by Elsevier Ltd.

1. Introduction

The United States Title 14 Code of Federal Regulations (CFR) defines *commercial air tours* as "flight[s] conducted for compensation or hire in an airplane or helicopter where a purpose of the flight is sightseeing (Federal Aviation Administration, 2012a)." Currently, air tours are conducted under CFR Part 91: General Aviation, Part 121: Domestic, Flag, and Supplemental Operations, and Part 135: Commuter and On Demand. Overall crash rates among Parts 91, 121, and 135 operators are 6.00, 0.30, and 1.06 crashes per 100,000 flight hours, respectively (BTS, 2012). These figures include both air tour and non-air tour operations, as industry-wide commercial air tour-specific crash rates have not been published. However, regional data from Hawaii indicate that commercial air tour crash rates are several times higher than those of other commercial flight operations (Haaland et al., 2009).

Li et al. (2008) developed a simple index for measuring the risk of fatal outcomes in aviation crashes. This score is called the "FIA Score" for the three risk factors it includes: ***F***ire, ***I***nstrument meteorological conditions, and being ***A***way from airport (Li et al., 2008). However, this score has not been validated as a risk measurement for fatal outcomes in commercial air tour crashes. This paper will describe the mortality burden of commercial air tour crashes, compare the characteristics of fatal and non-fatal crashes, and identify risk factors associated with fatal air tour crashes in the United States from 2000 through 2011. It will also assess the validity of the FIA Score in measuring the risk of fatality in commercial air tour crashes that occurred from January 2000 through December 2011.

* Corresponding author at: 615 N. Wolfe Street, Rm. 5515, Baltimore, Maryland, 21205, United States. Tel.: +1 410 614 3959.
E-mail address: sballard@jhsph.edu (S. B. Ballard).

0001-4575/$ – see front matter. Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.aap.2013.03.028

2. Methods

2.1. Data source and primary outcome measures

Study data came from the National Transportation Safety Board (NTSB) aviation crash surveillance system (NTSB, 2012). The Board is a Congressionally charged independent Federal agency that investigates all civil aviation crashes in the United States, as well as serious mishaps that occur in other transportation modes, such as highway, railway, marine, and pipeline (NTSB, 2012). For each serious transportation crash, the Board undertakes an investigation to determine the probable cause of the event, and it makes recommendations to prevent crashes in the future (NTSB, 2012). To ensure the objectivity of its reporting and recommendations, the Board has neither regulatory nor enforcement powers (NTSB, 2012).

An aviation "accident" is defined by the Board as "an occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight and all such persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage (NTSB, 2012)." "Death" is defined as a fatality occurring within 30 days of the crash (NTSB, 2012). "Serious injury" is defined as an injury that requires 48 h or more of hospitalization, causes a bony fracture (excluding simple fractures of the fingers, nose, or toes), damages any internal organ, results in severe hemorrhage, muscle, tendon, or nerve damage, causes second- or third-degree burns, or involves any burns affecting >5% of the body's surface (NTSB, 2012). "Substantial damage" is aircraft damage or failure requiring major component repair or replacement under normal circumstances (NTSB, 2012).

When an aviation crash occurs, the Board dispatches a field investigation team from its regional offices (NTSB, 2012). Crash data are collected and recorded on NTSB Form 6120.4, the Factual Report. The Factual Report details over 200 items pertaining to the crash circumstances, aircraft, and pilot (NTSB, 2012). Federal Aviation Administration Order 8020.11B contains detailed procedures that ensure the quality and technical depth of notification, investigation, and reporting procedures for aviation crashes (NTSB, 2012).

Commercial air tour crashes were identified by querying the National Transportation Safety Board's online Aviation Accident Database with the keywords "sightseeing," "sight seeing," "air tour," "airtour," and "sight." Two hundred fifty crashes were initially identified with this text search. Excluded from the study were 48 balloon crashes and 6 glider crashes because they did not meet the Federal Aviation Administration's "airplane or helicopter" requirement for classification as a commercial air tour flight. Additionally, 44 helicopter and airplane crashes were excluded because they were not conducted "for compensation or hire." A second query, performed using the National Safety Transportation Board's mainframe, returned the same cases. Unlike the publicly available query site, the internal database allows users to search for commercial air tours using a specific form-field search. Two reviewers read the probable cause and factual reports for each crash to ensure the commercial air tour crash definition was met. The 152 crashes that met the case definition were included in the analysis. The study was based on publicly available records and was exempt from review by the Johns Hopkins Bloomberg School of Public Health's institutional review board.

Total annual flight hours for Part 135 operators, Part 135 air tour operators, and Part 91 sightseeing flights were obtained from the Federal Aviation's General Aviation and Air Taxi Activity Survey for the period 2000 through 2010 (Federal Aviation Administration, 2010). The Federal Aviation Administration does not collect air tour-specific flight hour data for Part 91 commercial air tour operations, so the total number of Part 91 sightseeing flight hours was used as the denominator for Part 91 crash rate calculations. Because not all Part 91 sightseeing flights are conducted for compensation or hire, using this number underestimates the industry-wide and Part 91 commercial air tour crash rates.

2.2. Statistical analysis

Crash rates per 100,000 flight hours were calculated using numerator data from the National Transportation Safety Board's Aviation Accident Database and denominator data from the Federal Aviation Administration's General Aviation and Air Taxi Activity Survey (NTSB, 2012; Federal Aviation Administration, 2010). Data from the 152 commercial air tour crashes that occurred between January 1, 2000, and December 31, 2011, were analyzed in Stata version 12.0 (College Station, TX) (StatsCorp, 2011).

Fisher's exact test was used to examine fatal and non-fatal crashes for differences in binary and categorical exposures. *t*-Tests were used to examine differences in continuous exposure measurements between fatal and non-fatal crashes. The statistical analysis for predictors of fatal crash outcomes was performed in several steps. First, the association of fatal outcomes with fire, meteorological conditions, off-airport location, loss of power, loss of flight controls, maintenance malfunction, pilot error, US region (Alaska, Hawaii, or the 48 continental states), aircraft type (helicopter versus airplane), CFR category (Part 91 versus Parts 121/135), total pilot flight hours, and time of day were examined using simple and multiple logistic regression. Total pilot flight time, time of day, aircraft type, and CFR category were considered confounders based on prior knowledge and used to adjust the regression models. Multiple linear regression analysis permitted calculation of the variance inflation factors, which were all below 2.0.

Model-wise deletion was used to exclude missing values during all analyses; cases that had missing values for any of the variables used in the analytic model were excluded. Nested models were compared using likelihood ratio tests and Aikaike's information criteria. When two nested models were compared using likelihood ratio tests, only the cases common to both regression models following model-wise deletion were used for comparison.

Models were selected based on a combination of knowledge of crash survivability, potential confounders, and the projected feasibility of model use in the emergency response setting. Potential interactions were evaluated with regression models. Model fit was assessed using Pearson's and Hosmer–Lemeshow goodness of fit tests and deviance measures. Model performance was examined using the area under receiver operating characteristic (ROC) curves.

2.3. Risk index construction

Using the methods described by Li et al. (2008), a composite score was constructed from the three well-recognized predictors of fatal outcomes from the regression model: fire, instrument meteorological conditions, and off-airport location. The score was constructed with the following two assumptions: (1) the risk factors each have the same magnitude of influence on fatal outcomes, and (2) the risk factors have an additive effect on fatal outcomes that does not vary by combination of predictors (Li et al., 2008). Giving the fire predictor various weights, resultant differences in model performance were used to test the first assumption (Li et al., 2008). The second assumption was tested by comparing the rates of fatal outcomes between different combinations of predictors (Li et al., 2008).

The composite score was calculated for each of the 152 crashes in the study. Computed values for sensitivity, specificity, and area under the ROC curves were used to assess the validity of the score in measuring fatal outcomes. The ROC curve depicts the composite