Statistical Evaluation of Causal Factors Associated with Astronaut Shoulder Injury in Space Suits

Allison P. Anderson; Dava J. Newman; Roy E. Welsch

INTRODUCTION: Shoulder injuries due to working inside the space suit are some of the most serious and debilitating injuries astronauts encounter. Space suit injuries occur primarily in the Neutral Buoyancy Laboratory (NBL) underwater training facility due to accumulated musculoskeletal stress. We quantitatively explored the underlying causal mechanisms of injury.

METHODS: Logistic regression was used to identify relevant space suit components, training environment variables, and anthropometric dimensions related to an increased propensity for space-suited injury. Two groups of subjects were analyzed: those whose reported shoulder incident is attributable to the NBL or working in the space suit, and those whose shoulder incidence began in active duty, meaning working in the suit could be a contributing factor.

RESULTS: For both groups, percent of training performed in the space suit planar hard upper torso (HUT) was the most important predictor variable for injury. Frequency of training and recovery between training were also significant metrics. The most relevant anthropometric dimensions were bideltoid breadth, expanded chest depth, and shoulder circumference. Finally, record of previous injury was found to be a relevant predictor for subsequent injury. The first statistical model correctly identifies 39% of injured subjects, while the second model correctly identifies 68% of injured subjects.

DISCUSSION: A review of the literature suggests this is the first work to quantitatively evaluate the hypothesized causal mechanisms of all space-suited shoulder injuries. Although limited in predictive capability, each of the identified variables can be monitored and modified operationally to reduce future impacts on an astronaut's health.

KEYWORDS: extravehicular activity, shoulder injury, statistics, space suit.

Anderson AP, Newman DJ, Welsch RE. *Statistical evaluation of causal factors associated with astronaut shoulder injury in space suits.* Aerosp Med Hum Perform. 2015; 86(7):606–613.

Astronaut injury resulting from tasks performed in the space suit is one of the most important issues to be resolved with the current NASA space suit, the Extravehicular Mobility Unit (EMU).[12,14] Shoulder issues are the most serious and debilitating injuries associated with extravehicular activity (EVA) training.[16] The most recent report shows the number of astronaut shoulder surgeries has risen to 23 procedures, 11 of which are directly attributable to working in the space suit.[11] The causal mechanisms of these shoulder injuries have not been previously quantitatively evaluated.

EVA training is primarily performed in the neutral buoyancy laboratory (NBL). The NBL is a 6.2 million gallon swimming pool facility at the NASA Johnson Space Center used to simulate the weightlessness of microgravity for high fidelity training. For each hour of planned on-orbit EVA, an astronaut may spend an average of 11.6 h in NBL training.[14] Additionally, even before being assigned to a mission, each astronaut candidate goes through skills and maintenance training to become familiar with working in the EMU. An astronaut may spend their entire career working intermittently in the NBL. As a result of this accumulated time in the EMU, it has been well documented that astronauts are experiencing shoulder injuries.[11,14,16]

There are several hypothesized causes of space suit related shoulder injuries in the training environment of the NBL. The first is the design of the space suit hard upper torso, or HUT, and its restriction of shoulder movement. The HUT is a hard fiberglass shell forming the central structural component of the EMU on which other suit pieces are mounted. In training,

From the Massachusetts Institute of Technology, Cambridge, MA.

This manuscript was received for review in November 2014. It was accepted for publication in March 2015.

Address correspondence to: Dava J. Newman, 77 Massachusetts Ave., 33-307, Cambridge, MA 02139; dnewman@mit.edu.

DOI: 10.3357/AMHP.4220.2015

there are two HUT styles astronauts select to wear, namely, the pivoted and planar HUTs, shown in **Fig. 1**. The older style is the pivoted HUT and has a bellows at the scye shoulder bearing, giving it greater mobility over the planar HUT. The planar HUT's rotating bearing attaches directly to the upper arm piece. A full description of provocative shoulder motion for EVA shoulder injury is found in Williams and Johnson.[16] It is suggested that when scapulothoracic motion is restricted, normal shoulder movement is prevented. To compensate, astronauts rely more heavily on the rotator cuff muscles, causing overuse and injury. Additionally, when the astronaut shifts inside the suit due to gravity, the shoulder may be further impinged, causing inflammation of the fluid filled bursa and tendons below the clavicle.[14,16] Currently, inverted (or head down with feet up upside down body posture) NBL training is performed, but in limited duration. Since shoulder injuries persist despite limited time spent in the inverted position, evidence is mounting that HUT design and shoulder movement are the dominant factors in injury.[11]

There are several training factors associated with the NBL that may contribute to propensity for astronaut injury. The viscosity of water adds resistance to movement, causing astronauts to incur additional metabolic costs.[6,7] Although made neutrally buoyant, the tools used by astronauts often have high inertia, causing musculoskeletal stress as astronauts generate torques, change their momentum, or move tools from site to site. Exacerbating the problem, time constraints in preparing for a mission may also prevent astronauts from fully recovering between trials in the NBL.[14–16] Over the course of a career, these musculoskeletal stresses may contribute to shoulder degeneration.

Finally, individual variability of each person due to body morphology or propensity for shoulder injury may be contributing factors.[11,14,16] Movement in the suit is limited and unnatural due to the space suit's inherent programming, or planes through which the suit is able to move due to the angle of rotational bearings,[2] and gas pressurization causing stiffness and rigidity.[5,8,13] Astronauts must learn to change their biomechanical movement strategies, rather than attempting to move as they do unsuited.[3] Additionally, how a person's body fits inside the suit is a critical factor to achieve optimal movement and the best performance, but no two people maintain the same body position relative to the suit due to natural variation.[1]

Previous work analyzing shoulder injury has focused primarily on tracking injury incidence in both orbital and training environments.[4,12,15] Recently, shoulder surgery data was matched with the crewmember's HUT selection to look for statistical correlations between HUT style (planar vs. pivoted) and frequency of space suited activity. This work found the planar HUT to be the most provocative for injury, but there is huge individual variability between those who get injured and those without injury.[11] Additional data, such as subject anthropometry and training history may provide further insight into shoulder injury, but has not been previously explored. Information regarding shoulder injury, anthropometry, and training in the NBL is spread among many different groups at NASA, each with variable reporting criteria. Investigating how these factors contribute to shoulder injury is critical to understanding and mitigating the problem.

The objective of this statistical analysis investigating astronaut shoulder injury is to explore the relationship between anthropometry, space suit HUT design, and training data as they impact shoulder injury. We hypothesized that each of these variables would be a predictive factor in identifying astronauts with a reported shoulder incident. Each of the hypotheses investigated a specific causal mechanism found in the literature associated with EVA shoulder injuries. Each factor was evaluated by statistical regression to determine which variables contributed to an increased propensity for astronaut injury.

Fig. 1. Pivoted and planar hard upper torso (HUT) styles. There are two HUT styles with their primary difference being the shoulder scye bearing to which the arm components attach. The pivoted HUT, shown on the left, has a pivoting bellows to allow for greater shoulder mobility than the planar HUT, shown on the right. Photo credit NASA.

METHODS

Database

The research effort made use of an extensive, new database which was compiled by NASA personnel at the Longitudinal Study on Astronaut Health (LSAH), including three components: anthropometric measurements, training record, and injury record. Each astronaut in the database was given a unique identifier and all data were made nonattributable. Due to the many resources compiled to create this comprehensive database, there is some variation in which subjects are found in each of the three sections. There are a total of 278 astronauts with information in at least 1 of the 3 sections. However, only 119 of the astronauts are common to all 3. The remaining astronauts have data in at least one of the remaining sections, summarized in **Table I**.

There are 16 anthropometric dimensions included in the database, focusing primarily on the upper body: height (stature), cervical height, mid-shoulder height (left and right), acromion height, arm reach (left and right), expanded chest depth, inter-acromion distance, chest breadth, bi-deltoid breadth, acromion radiale length (left and right), lower arm length (left and right), and shoulder circumference. There are 180 astronauts with reported anthropometry dimensions. This data was collected in sizing astronauts to determine best suit fit. Anthropometric information is known to be normally distributed and highly correlated within a population.[9] However, this assumption may not be met with small sample sizes and without separating subjects by gender. Each dimension was checked for outliers and seven data points were removed. These outliers were attributable to errors in entering the data, as confirmed by NASA personnel. The data set was normally distributed using the Kruskal-Wallace test for normality ($P > 0.05$). Of the possible 120 correlation coefficients, 116 had $P < 0.05$. Therefore, the anthropometric data was treated as approximately univariate normal and highly correlated, as expected.

The astronaut training record contains five different sets of information: training day, either the actual or estimated time in the space suit, whether the subject was wearing either the planar or pivoted HUT, and the size of his or her HUT on a given training day. The training day variable begins with the subject's first time in the NBL, and continues sequentially over the duration of his or her career. For some subjects, HUT size and training time were not estimated or recorded, and therefore these variables were not included in the regression models. Each training incident, however, does include the HUT type (planar vs. pivoted) worn by the astronaut. There are 224 astronauts in the training record. These astronauts were on active duty from 1981 to 2012. There are 12,170 training events recorded. The training record data was aggregated into five different dimensions, summarized in **Table II**. Each dimension is a proxy variable to capture a specific aspect of the training history that may or may not play a role in shoulder injuries. None of the dimensions is normally distributed.

The injury record includes every shoulder incident reported by an astronaut, whether it occurred preselection, in active duty, or during retirement. Shoulder incidents are recorded by the date of the report (although this may not correspond to the date of the injury), date of surgery if one occurred, whether a relationship to training in the water immersion training facility (WITF) or NBL was noted, precursory events, diagnosis, and the subsequent treatment. Although there are 196 astronauts with reported shoulder incidents, only a small subset is relevant to the research questions explored in this research effort. Incidents were evaluated to divide subjects into four groups. Note that although referred to as an "injury," not all shoulder incidents categorized from the database are considered a medical injury. However, any reported incident may be relevant and is considered in these models. The four groups are: 1) those whose injuries are reported to be directly attributable to the space suit or training environment; 2) those whose shoulder pathologies began during active duty so the suit or training environment may be a contributing factor; 3) those with shoulder pathologies beginning either prior to selection or after retirement, indicating shoulder injuries may be a result of normal shoulder deterioration with the suit/training environment as a potentially contributing factor; and 4) those whose injuries were reported as directly attributable to something other than working in the NBL or space suit environment (for example, an injury caused by a bicycle accident).

Table I. Common Subjects Between Three Components of the Database.

DATABASE	ANTHROPOMETRY	TRAINING	INJURY
Anthropometry	180		
Training	180	224	
Injury	119	142	196
Total	278		
Common	119		

Each database was compiled from a different resource and, therefore, has some variability in the subjects presented. There are 278 subjects, 119 of whom were common to all 3 databases.

Two groups of subjects were evaluated. The first group of subjects are those whose injuries are directly attributable to the space suit or training environment, henceforth referred to as the NBL group. There were 35 subjects in this category. The second group of subjects are those whose pathologies began in active duty and are combined with the NBL subjects for a second statistical analysis. This group is referred to henceforth as the Active group. There were 62 subjects in this category.

There is a great deal of variability in the recording of astronaut shoulder incidents. Historically, there was a tendency to under-report or delay reporting, so as not to affect the astronaut's flight status and career. Additionally, there was no standard method by which information was recorded, so details vary with each report. With the recent attention to shoulder injuries, many of these issues are being resolved and future data will be recorded more systematically. However, when dealing with historical information, this is an important limitation to keep in mind.

Statistical Methods

A logistic regression was chosen to analyze the relationship between training data, anthropometry, space suit components, and shoulder incidents. Logistic regression does not require any underlying assumptions about the distribution of the predictor variables. Finally, it is used to regress against a binomial response, in this case either injured or uninjured. The equation for logistic regression is:

$$Y_i = 1/\left(1 + e^{-\boldsymbol{X}'\beta}\right) \qquad \text{Eq. 1}$$

where Y_i is the logistic response function, $\boldsymbol{X}$ is the matrix of observations for each explanatory variable, and β is the matrix of fit coefficients. When expanded for each explanatory variable:

$$X'\beta = \beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n \qquad \text{Eq. 2}$$

To determine a useful set of variables to include in the regression models, the following method was used. This analysis was performed using MATLAB software (MathWorks, Natick, MA).

The data used for the regression was compiled from the information in the previously described shoulder injury database. There are relatively few data points and a large number of potential predictor variables. Because this study focuses on shoulder injury, lower arm dimensions were excluded. In addition to the 119 subjects common to all 3 sections of the database, there are an additional 61 subjects both with anthropometry and training information in the database, but who never reported a shoulder incident. These subjects were included in the model as uninjured, bringing the total number of subjects evaluated to 180 astronauts.

Anthropometric data was centered about the mean and normalized by SD (σ). The training data was also scaled and centered, but since the variables are not normally distributed, the median and median absolute deviation (MAD) were used. Three variables, the total number of incidences, training frequency, and recovery, were exponential in nature and were log-transformed to improve the fit of the model. For the NBL injured subjects, an additional predictor variable was included for whether or not the subject had been previously injured.

Given the large number of predictor variables, bootstrapping was used to identify the most relevant factors.[10] A model fit to the entire data set may lose some of the subtly relevant predictor variables in favor of a more parsimonious model. This is particularly true for the NBL injured group of subjects where an injury is an infrequent occurrence. A 500 iteration bootstrap was fit to a 50-50 data split where every injured subject was used each time to build a model, but the uninjured subjects were randomly selected using resampling with replacement to equal the number of injured subjects. Forward stepwise logistic regression was used to fit the model with the decision criteria for inclusion being to minimize the Akaike's Information Criterion (AIC) statistic. For each of the 500 models built, the relevant predictor variables were logged. For the NBL subjects, variables appearing in 10% of the models were considered for inclusion in the final model. For the Active subjects, variables appearing in 30% of the models were considered for inclusion. These cut-off values divided variables with substantially more logged occurrences from the rest.

Using the variables identified in the bootstrap, a model was fit to the entire dataset. The variables were checked for multicollinearity using the variance inflation factor (VIF). A stratified fivefold cross validation was performed using the final predictor variables to determine the model's fit to "unseen" data. The 180 subjects were randomly divided into 5 equal sections, preserving the global incidence rate of injured and uninjured subjects. A model was built on 80% of the data and tested using the remaining 20% of the data. Nominally, the model was fit with the cut-off value of 0.5, above which a subject is categorized as injured. The cut-off value was shifted to improve prediction rates of injured subjects, trading off correctly identifying injured subjects as more important than miscategorizing uninjured subjects given the detriment associated with misclassification. The percent of correct predictions, percent of correct negative predictions, and percent of correct positive predictions were logged for each of the five models. To reduce the effects of randomness, cross validation was performed 50 times and mean and SD of predictive capability were calculated over all 250 trials.

RESULTS

The final logistic regression model fit to the NBL identified category is shown in **Table III**. There were 35 astronauts considered injured in this model, and 145 considered uninjured. There were three relevant predictor variables related to training: the percent time in the planar HUT, training frequency, and recovery. Two anthropometric dimensions were found to be important predictors for injury: expanded chest depth and bideltoid breadth. Finally, history of a previous injury was found to be a relevant predictor. This model has a log-likelihood overall model fit P-value = 0.003. Despite strong correlations in the

Table II. Aggregated Training Information Used as Proxy Variables for Statistical Analysis.

PREDICTOR VARIABLE	DESCRIPTION	FORMULA	VARIABLES
Total Incidence	Total number of training events	$\sum_{i=1}^{n} k_i$	*i*: observation *n*: total observations *k*: training event
Percent in Planar HUT	Percent of training events performed in the Planar HUT	$100 * \sum_{i=1}^{n} k(p)_i / \sum_{i=1}^{n} k_i$	*p*: planar event
Longevity	Total number of active duty days	$s_n - s_1$	*s*: training date
Training Frequency	Average number of active duty days per training event	$(s_n - s_1) / \sum_{i=1}^{n} k_i$	
Recovery Metric	Measure of how much recovery subjects received between training runs; Sum of the inverse of days between training	$\sum_{i=1}^{n-1} \frac{1}{s_{i+1} - s_i}$	

Five variables were created by aggregating the training data for each subject. Variables are not normally distributed.

Table III. Model Fit to Subjects Whose Incident Was Reported as a Result of Working in the NBL.

	COEFFICIENT	VARIABLE	WALD STATISTIC	*P*-VALUE
β0	−1.79	Constant	−6.59	-
β1	1.06	Percent in planar HUT	3.44	0.0006*
β2	0.073	Training frequency	0.47	0.64
β3	0.42	Recovery metric	2.09	0.037*
β4	−0.33	Expanded chest depth	−1.15	0.25
β5	0.19	Bideltoid breadth	0.7	0.48
β6	0.98	Previous injury	1.66	0.1

Six predictor variables were found to be important for identifying injury: three related to training, two anthropometric dimensions, and record of previous injury.
* Indicates the predictor variable is significant in the model, $P < 0.05$.

anthropometric data, each variable had a VIF less than 2. The area under the receiver operating characteristic (ROC) curve is 0.73, shown in **Fig. 2**, and the Hosmer-Lemeshow test for fit is not significant ($P = 0.84$). Each of these metrics indicates the model fits the data well. When evaluated in cross-validation, using a cut-off value of 0.3, the model had a 69% overall accuracy rate and correctly identified 39% of injured subjects as injured (SD, $\sigma = 9\%$).

The same methodology was used for subjects in the Active category combined with the NBL subjects. There were 75 astronauts considered injured in this second model and 105 considered uninjured. The record of previous injury was not included in this analysis since it is a confounding variable with the way Active injured subjects were categorized. The final model is shown in **Table IV**. A total of five predictor variables were found to be important for identifying subjects as injured. There are three variables related to training: percent incidences in planar HUT, frequency of training, and recovery. Two relevant anthropometric predictor variables were identified: expanded chest depth and shoulder circumference. This model has a log-likelihood overall model fit with P-value $= 0.003$. Again, despite strong correlations in the anthropometric data, each variable in the model had a VIF less than 2. The area under the ROC curve is 0.67, shown in Fig. 2, and the Hosmer-Lemeshow test was not significant ($P = 0.89$), also indicating this model is a good fit to the data. The cross-validated correct prediction rate is 57% with a cut-off value of 0.4. The correct prediction of injured subjects is 68% ($\sigma = 10\%$).

Fig. 3 shows the correlation between each variable used in both the NBL and Active models, excluding the categorical variable 'previous injury'. Histograms are also given for each. In addition to the models presented here, additional models were evaluated for both groups using only anthropometric information and only training information. For each, models built with only training information were found to be significant. However, anthropometry alone did not produce a significant result. For all cases, cross-validated performance was poor, therefore only models using anthropometry in conjunction with training information and record of previous injury (NBL subjects only) were considered, which achieved a better overall model fit and improved predictive performance.

DISCUSSION

For both groups of injured subjects, the NBL and Active groups, a logistic regression model was calculated with a statistically good fit to the data. The two models use similar predictor variables.

For both models, percent of training incidences in the planar HUT is a highly significant factor. HUT type is consistently the best predictor of injury as well as the most frequently identified variable in bootstrapping, thus confirming our hypothesis that space suit training variables in the planar HUT, rather than training in the pivoted HUT, will be a predictive factor in identifying astronauts with a reported shoulder incident. Although HUT style has been reported as a major cause based on anecdotal evidence,[14,16] it has not been until recently that this causal mechanism has been quantitatively evaluated.[11] The findings of Scheuring and McCollouch are corroborated by these new results, but expands upon them to include additional relevant factors not previously explored. This work also includes other shoulder incidents not evaluated by Scheuring et al., which, although not defined as medical injuries, have had negative impact on crew comfort and health, as well as impacting an astronaut's operational availability. It has long been asserted that training in the planar HUT is the most relevant factor, and these results support this assertion for both groups of injured subjects.

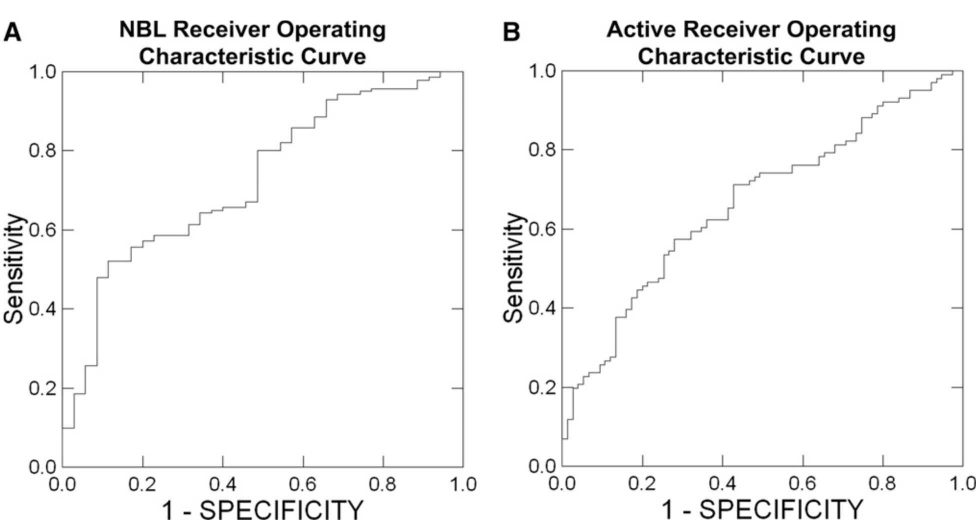

Fig. 2. Receiver operating characteristic curve. A) For the NBL model. The area under the curve is 0.73. B) For the Active model. The area under the curve is 0.67.

Table IV. Model Fit to Subjects Whose Incident Was Reported During Active Duty and While Working in the NBL.

	COEFFICIENT	VARIABLE	WALD STATISTIC	*P*-VALUE
β0	−0.37	Constant	−2.2	-
β1	0.55	Percent in planar HUT	3.0	0.003*
β2	0.023	Training frequency	1.89	0.06
β3	0.37	Recovery metric	2.36	0.02*
β4	−0.31	Expanded chest depth	−1.36	0.17
β5	0.5	Shoulder circumference	2.34	0.02*

Five predictor variables were found to be important for identifying injury: three related to training and two anthropometric dimensions.
* Indicates the predictor variable is significant in the model, $P < 0.05$.

The same training variables, frequency and recovery, were included in both models. As described in Table II, frequency is a measure of how often the astronaut trains over his or her career, while recovery is a measure of the concentration of consecutive runs. These variables confirm the hypothesis that operational training variables will be predictive factors in identifying astronauts with a reported shoulder incident. This supports the conclusions reached[16] by regarding the import of the training environment as a contributory factor, but this is the first quantitative assessment of the impacts of training frequency and recovery. Although frequency was not significant in the NBL model by the Wald criteria, which evaluates whether the factor contributes significantly to the model, it was included to improve predictive power. Note that both the recovery metric and the total number of training incidents are nearly equivalent predictors of injury because they are strongly correlated (value of 0.93). Recovery was chosen over total training incidence because it improved the correct prediction of injured subjects as compared to the former, and had a much lower VIF (for total training incidence VIF > 10). However, total training incidence should be considered as an important variable for future modeling work as more data is collected. These factors show that an astronaut who trains frequently will have a higher propensity for injury, in addition to whether or not those training runs are over a concentrated time period. Although these may seem like confounding factors, as seen in Fig. 3, the correlation between these variables is low (0.22). We recommend recovery and training frequency be altered operationally to improve crew health and safety. Previous work has suggested changing body posture orientation during training and providing additional assistance to astronauts would reduce injury,[16] but since injuries have persisted, our results suggest additional alterations to training to reduce frequency and increase recovery should be implemented.

Additionally, anthropometric variables were found to be relevant for both models, confirming the hypothesis that anthropometric dimensions will be a predictive factor in identifying astronauts with a reported shoulder incident. Bideltoid breadth, expanded chest depth, and shoulder circumference were found to be the anthropometric dimensions most strongly related to injury. These dimensions affect how the person fits inside the HUT and, therefore, how their motions are achieved. These particular body dimensions should be the focus for future space suit design studies and to ensure astronauts are working in the HUT that fits them best.[1,3] Expanded chest depth was shown in both models to provide explanatory power. As a variable with a negative coefficient, a decrease in expanded chest depth will increase the odds of being injured. It has been proposed in the literature that smaller subjects who must work inside a HUT that is too large for them may have additional problems than those described previously to articulate the suit due to the lateral shifting of the scye bearing, potentially leading to injury.[16] Although our results cannot support or refute this claim, we do find support that smaller expanded chest depth increases propensity for injury.

One additional anthropometric dimension was chosen for each model, but the variable is different for each group of subjects.

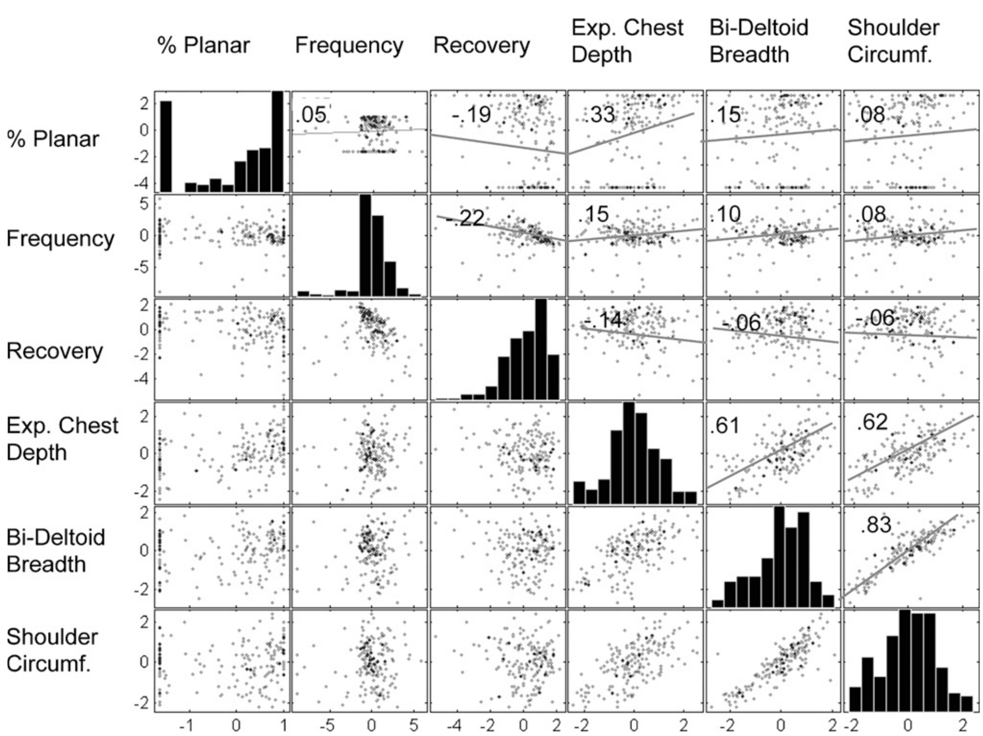

Fig. 3. Correlation matrix and histogram of variables included in the NBL and Active injury models. Histogram of each variable is given on the diagonal axis. The pairwise correlation is given in each row/column pair with the correlation coefficient. The highest correlation in each model is between two anthropometric dimensions with values of 0.61 and 0.62, respectively. The highest correlation is between bideltoid breadth and shoulder circumference, not used simultaneously in either model. Categorical variable of previous injury not shown.

For NBL subjects, bideltoid breadth was an important predictor, whereas for the Active subjects, shoulder circumference was a strong predictor and was statistically significant. Although, as seen in Fig. 3, these variables are correlated with one another, they were poor predictors when used in the opposite model (i.e., replacing bideltoid breadth for shoulder circumference in the NBL model does not give a well fit model). For each model respectively, if the bideltoid breadth or shoulder circumference increases, the odds of getting injured also increases, as opposed to the smaller expanded chest depth. This variable seems to be identifying injured subjects on the larger spectrum who are potentially fitting more tightly into their HUT, not allowing for normal body movement. Subjects with less clearance for scapular thoracic motion may not be able to move as unsuited, leading to shoulder injuries.[16] Although this work cannot confirm or support this claim, it does indicate it is an interesting area for future inquiry.

Finally, the hypothesis investigating the record of previous injury as a predictive factor in identifying astronauts with an additional shoulder incident is confirmed. Although there are many astronauts with previous shoulder injuries without subsequent problems, and many without previous injury as evidenced by the Active group of subjects, our results show that for NBL subjects previous injury is a strong predictor variable. This may be due to normal shoulder deterioration, and personnel at the Johnson Space Center are currently working an age-matched incident rate against which to compare astronaut shoulder injury incidence rates. Regardless, this information allows flight surgeons and astronaut strength and conditioning personnel to identify higher risk astronauts to ensure they are properly trained and healthy before entering the NBL training environment.

Although these models fit the data by objective measures, there is an inherent optimism in that their performance is evaluated against the data from which it was fit. Cross-validation allows us to understand how the model performs on "new" data that was not seen when the model was built. Ideally, the model would separate the injured from the uninjured with no type I (false positive) or type II (false negative) errors. For the purposes of predicting astronaut shoulder injury, type I error is favored due to the consequences of misidentifying a subject who will be injured at the cost of crew health and safety and mission success. The NBL model with a cut-off value of 0.3 has a reasonably high overall accuracy rate of 69%. However, it only correctly identified 39% of injured subjects. Although predicting any injured subject correctly is an improvement over the current state, it is desirable for this rate to be higher. The overall prediction rate was sacrificed to increase the subjects who could be identified as injured by shifting the cut-off value. Only 19% of the subjects in the NBL data are injured and, therefore, it has a tendency to predict subjects as uninjured. However, for the Active model, the incident rate is higher: 42% or 75 injured astronauts. Here, the overall prediction is 57% with a cut-off value of 0.4, but the correct prediction of injured subjects is 68%. The ability to identify injured subjects is greatly improved in this second model, at the cost of misclassifying subjects who were uninjured in reality. Note that shifting the classification cut-off value back to the original 0.5 does not improve the overall prediction rate, but rather moves subjects from type I to type II error, which is undesirable. Both the NBL and Active models are able to identify some subjects as injured, but their performance is not as strong as desired. The models cannot fully separate subjects, but rather pushes injured subjects closer to the front where we can identify them as injured. This is similarly reflected when the residuals are evaluated, where for the NBL model, large deviation from normality for the injured subjects and deviation at both tails for the Active model is seen. This indicates our statistical models are missing critical information to better identify injured and uninjured subjects properly. However, given the current data set, the models presented herein provide the most utility to date.

This research effort provides a framework for identifying relevant predictor variables related to space suit injury given the small number of data points and large number of predictor variables with different distributions. Additional data would improve the statistical analysis. In the past, injuries have not been reported in a uniform manner. Finer detail would improve the categorization of injured and uninjured subjects, or even allow detailed analysis to be performed on specific injuries. NASA's current efforts to centralize injury reporting may address this issue in the future. Higher fidelity 3D biomechanical models would allow clearance analysis to be performed, giving a more precise understanding of how the astronaut fits inside the space suit. Future work includes evaluating more complicated statistical modeling paradigms. Several techniques have already been explored. Advanced methods include linear discriminant analysis, partial least squares regression, principal component analysis, decision trees, and random forests.

The results presented herein address the current gap in our understanding of the causal mechanisms of astronaut injury inside the space suit. An astronaut's favored HUT style, recovery time, training frequency, upper body anthropometric dimensions, and record of previous injury were each found to be contributing variables in identifying astronauts with a reported shoulder incident. The statistical models cannot identify all astronauts that may be injured as a result of working inside the space suit due to individual variability and the data available. These models, however, represent an improvement over the current understanding of shoulder injury mechanisms. This work provides quantitative confirmation of many of the assertions made previously by flight doctors, trainers, space suit designers, and astronauts about relevant factors of injury inside the EMU and NBL training environments. Future human space exploration will continue to require EVA, potentially leading to higher injury incidence if the space suit system is not enhanced to find long-term, healthy solutions to EVA injury.

ACKNOWLEDGMENTS

The authors would like to thank our colleagues at NASA and the LSAH for their support and collaboration. This project was funded through NASA Grant NNX12AC09G, "Spacesuit Trauma Countermeasure System for Intravehicular and Extravehicular Activities." Additional support was provided by the National Science Foundation Graduate Research Fellowship Program, Singapore-MIT Alliance for Computation Systems and Biology, and the MIT Center for Computational Research in Economics and Management Science.

Authors and affiliations: Allison P. Anderson, Ph.D., M.S., Geisel School of Medicine, Dartmouth College, Lebanon, NH; Dava J. Newman, Ph.D., M.S., Massachusetts Institute of Technology, Cambridge, MA; and Roy E. Welsch, Ph.D., M.S., Sloan School of Management, MIT, Cambridge, MA.

REFERENCES

1. Benson E, Rajulu S. Complexity of sizing for space suit applications. In: Duffy VG, editor. Digital human modeling. Berlin Heidelberg: Springer-Verlag; 2009:599–607.
2. Cowley M, Margerum S, Harvill L, Rajulu S. Model for predicting the performance of planetary suit hip bearing designs. In: Duffy VG, editor. Advances in human factors and ergonomics. Boca Raton (FL): CRC Press; 2012.
3. Gast M, Moore S. A glimpse from the inside of a space suit: whats it really like to train for an EVA? Acta Astronaut. 2011; 68(1-2):316-325.
4. Hochstein J. Astronaut total injury database and finger/hand injuries during EVA training and tasks.Illkirch-Graffenstaden (France): International Space University; 2008.
5. Holschuh B, Waldie J, Hoffman J, Newman D. Characterization of structural, volume, and pressure components to space suit joint rigidity. International Conference on Environmental Systems. Savannah (GA): Society of Automotive Engineers; 2009.
6. Newman D. Human locomotion and energetics in simulated partial gravity. Cambridge (MA): Massachusetts Institute of Technology; 1992.
7. Newman DJ, Alexander HL. Human locomotion and workload for simulated lunar and Martian environments. Acta Astronaut. 1993; 29(8): 613–620.
8. Parry DG, Curry LR Jr, Hanson DB, Towle GB. A study of techniques and equipment for the evaluation of extravehicular protective garments. Dayton, OH: Hamilton Standard; 1966.
9. Proctor R, Van Zandt T. Anthropometrics and workspace design. In: Proctor RW, Van Zandt T. Human factors in simple and complex systems, 2nd ed. Boca Raton (FL): CRC Press; 2008:433-466.
10. Rice J. Mathematical statistics and data analysis, 3rd ed. Belmont (CA): Brooks/Cole; 2007.
11. Scheuring R, McCullouch P, VanBaalen M, Minard C, Watson R, et al. Shoulder injuries in U.S. astronauts related to EVA suit design. [Abstract.] Aviat Space Environ Med. 2012; 83(3):313.
12. Scheuring RA, Mathers CH, Jones JA, Wear ML. Musculoskeletal injuries and minor trauma in space: incidence and injury mechanisms in U.S. astronauts. Aviat Space Environ Med. 2009; 80(2):117–124.
13. Schmidt P. An investigation of space suit mobility with applications to EVA operations. Cambridge (MA): Massachusetts Institute of Technology; 2001.
14. Strauss S. Extravehicular Mobility Unit training suit symptom study report. Houston (TX): Johnson Space Center; 2004.
15. Strauss S, Krog RL, Feiveson AH. Extravehicular mobility unit training and astronaut injuries. Aviat Space Environ Med. 2005; 76(5): 469–474.
16. Williams DR, Johnson BJ. EMU shoulder injury Tiger Team report. Houston (TX): NASA Johnson Space Center; 2003:104.