MILITARY MEDICINE, 176, 4:431, 2011

# Predicting Success in the Tactical Air Combat Party Training Pipeline

John Kalns, PhD*; Jonathan Baskin, MS*; Andrew Reinert, MS*; Darren Michael, PhD*; Adrienne Santos, MS*; Sheena Daugherty, BS*; Col James K. Wright, SFS, USAF MC (Ret.)†

**ABSTRACT** Objectives*:* To develop a statistical model that predicts the likelihood of success or failure of military training candidates using tests administered before initial skill training as inputs. Methods: Data were acquired from candidates before the start of U.S. Air Force Tactical Air Control Party training, including (1) demographic, (2) psychological composition evaluated using Emotional Quotient Inventory, (3) physical performance capability, (4) a physical activity questionnaire, and (5) salivary fatigue biomarker index. A total of 126 candidates were tracked until they either passed or failed the training, and a total of 55 variables were used as inputs for creation of the model. Results: Clustering analysis of the data revealed that only 4 of 55 variables were useful for predicting success or failure. The variables in the order of their importance are as follows: run time, number of miles run per week in the last year, level of salivary fatigue biomarker, and height. Conclusions: The results suggest that simple testing methods can identify candidates at high risk of failure.

## INTRODUCTION

Battlefield airmen comprises 4 groups of elite Air Force personnel who are responsible for specialized ground-support duties ranging from rescue and extraction of personnel from threatening environments to coordinating Special Operations and air strikes.[1,2] One such group, Tactical Air Control Party (TACP) specialists, consists of 2-member teams assigned to Army combat units around the world. These teams direct close air-support firepower toward enemy targets on the ground and advise Army combat commanders on the use of Air Force air power.[3] Training of TACP specialists requires approximately 1 year and an investment of $30,000 per airman. Unfortunately, the Air Force has not been able to meet its goals in producing TACP specialists in recent years and the problem has intensified during the last 2 years. For the TACP training pipeline, hundreds of training days are wasted each year because of failure to perform on the part of the trainee. In the 4 TACP classes comprising this study, approximately 50% of the trainees failed to complete the training successfully. Improved methods for selection of candidates at the start of the training would result in a higher training success rate and an increased manning in the TACP career field.

Using demographic, psychosocial, and physical performance and biomarker data sets, we have developed a mathematical relationship that predicts who is likely to succeed or fail in TACP training. Virtually all the data, including human performance biomarker data, are collected at the outset of the training cycle. Thus, it is possible that candidates at high risk of failure can be identified at the outset of training and support measures brought to bear, leading ultimately to success. Alternatively, assessment could be made before arriving for TACP training to select basic trainees having the highest probability of success during the initial TACP training cycle. Ultimately, the approach could prove useful at increasing pass rates through undermanned career fields without compromising the quality of graduates.

*Hyperion Biotechnology, 13302 Langtry Street, San Antonio, TX 78248.

†Former 720th Special Tactics Group Surgeon, 223 Cody Avenue, Hurlburt Field, FL 32544.

## MATERIALS AND METHODS

### Study Participants

The study was conducted with approval of the Wright-Patterson Air Force Base Institutional Review Board. All subjects provided written and verbal consent before the start of data collection. Subjects were recruited from four TACP classes during the period of August 2008 to February 2009. Subjects were assigned a number and all collected data were associated with the subject number to protect trainee confidentiality.

### Data and Sample Collection

The following data was collected after informed consent was obtained: (1) demographic questionnaire, (2) Bar-On Emotional Quotient Inventory (EQI) test,[4] (3) Armed Services Aptitude Battery (ASVAB), (4) physical fitness test, and (5) saliva sample. All data were collected before the start of the training classes. In some instances, subjects who initially failed to complete a training class were returned to the start of the training. In these instances, the physical fitness test and saliva samples used in the analysis were taken immediately before the "restart" of training.

### Demographic Questionnaire

The questions posed and type of responses allowed are shown in Table I. Questions with written responses were not included in subsequent analysis. Integer values were assigned to responses in the order shown in the table. For example, for question 21, a response of none was assigned a value of "0,"

**TABLE I.** Demographic Questionnaire

| Question | Responses |
|---|---|
| 1. Did You Play Sports During High School? | Yes or No |
| 2. How Many Miles Run per Week Did You Run During the 3 Months Before BMT? | Numerical |
| 3. How Many Hours per Week Do You Spend Strength Training in the 3 Months Before BMT? | Numerical |
| 4. What is the Average Number of Miles per Week that You Ran in the 12-month Period Before BMT? | Numerical |
| 5. What is the Average Number of Hours You Spend in Strength Training per Week in the 12-month Period Before BMT? | Numerical |
| 6. What Parts of Body that You Train Routinely? | Neck, Shoulders, Chest, Back, Abdomen, Arms, Legs |
| 7. How Do You Do Strength Training? | (1) Same Exercises and Use Heavy Weight, (2) Same Exercises and Use Moderate Weight, (3) Seldom Do the Same Exercises, but Use Heavy Weight, (4) Seldom Do the Same Exercises but Use Moderate Weight, (5) Others |
| 8. What Level of Expertise Do You Have With Hunting, Fishing, Camping, Hiking, Scouting, Water? | (1) Expert, (2) Somewhat Skilled, (3) A Little Experience, (4) No Experience |
| 9. Did You Grow Up in a Rural or Urban Environment During High School? | (1) Farm, Ranch, or in the Country, (2) Small Town <10,000, (3) Small City 10,000–100,000, (4) City >100,000, (5) Moved Around a Lot |
| 10. Do you have experience being in a position of leadership? | Yes or No |
| 11. On Average, in the Last 6 Months, How Many Hours Do You Spend per Week Playing Video Games? | (1) None, (2) 1–5 hours, (3) 5–15 hours, (4) >15 hours |
| 12. Do You Have Family in the Military? | Yes or No |
| 13. Did a Military Member Influence Your Decision to Join the Air Force? | Yes or No |
| 14. What is Your Highest Educational Level? | (1) High School, (2) 1 year in College, (3) 1–2 years in College, (4) 2–4 years in College, (5) >4 years in college |
| 15. How Confident are You That You Will Complete TACP Training? | (1) Very Confident, (2) Confident, (3) Somewhat Confident, (4) Not Confident, (5) Not Sure |
| 16. What Is Your Level of Understanding of the Training You Will Receive in TACP Training? | (1) Very Understanding, (2) Understanding, (3) Somewhat Understanding, (4) Don't Understand, (5) Not Sure |
| 17. Why Did You Choose the Air Force | Written Response |
| 18. How Would You Rate the Level of Training You Have Received Based on Your Expectations? | (1) Easier, (2) About the Same, (3) More Difficult, (4) Much More Difficult |
| 19. Why Did You Choose TACP? | Written Response |
| 20. Over the Past Year, How Many Days Were You Unable to Work or Perform Your Usual Activities Because of Injury? | (1) 0, (2) 1, (3) 2–5, (4) 6–10, (5) 11–15, (6) 16–20, (7) >20 |
| 21. Over the Past Year, How Many Days Were You Unable to Work Or Perform Your Usual Activities Because of Illness? | (1) 0, (2) 1, (3) 2–5, (4) 6–10, (5) 11–15, (6) 16–20, (7) >20 |

a response of one was assigned the value "1," a response of 2 to 5 was assigned a value of "3," and so on. Weight and height were measured by trainers. Age was calculated from the subject records.

### Emotional Quotient Measurement (Bar-On Emotional Quotient Inventory)

The Bar-On EQI is a 133-question self-report measure of emotionally and socially intelligent behavior that provides an estimate of emotional-social intelligence. Analysis is conducted using a proprietary algorithm that reduces the test questions to numerical outputs in 5 composite scales and 15 subscales. The composite scales and associated subscales are as follows: (1) Intrapersonal Scales: self-regard, emotional self-awareness, assertiveness, independence, self-actualization; (2) Interpersonal Scales: empathy, social responsibility, interpersonal relationship; (3) Adaptability Scales: reality testing, flexibility, problem solving; (4) Stress Management Scales: stress tolerance, impulse control; and (5) General Mood Scales: optimism, happiness. For the purpose of this study, the scores from subscales were used as inputs into predictive models.

### Armed Services Vocational Aptitude Battery

The ASVAB is a 200-question test. The ASVAB measures strengths, weaknesses, and potential for future success by testing the following areas: general science, arithmetic reasoning, word knowledge, paragraph comprehension, auto and shop information, mathematics knowledge, mechanical comprehension, and electronics. The composite scores in the following areas of verbal, mathematical, and academic ability were used as inputs into mathematical models.

### Physical Fitness Testing

The physical fitness test includes (1) time in minutes required to run 1.5 miles, (2) number of sit-ups performed in 1 minute, (3) number of push-ups performed in 1 minute, and (4) the number of pull-ups performed in 1 minute.

### Fatigue Biomarker Measurement

The relative abundance of a specific heptapeptide, ESPSLIA, in saliva has previously been shown to correlate with exertional fatigue (Unpublished result). When measured against the concentration of a second salivary heptapeptide, GGHPPPP, a "fatigue biomarker index" can be calculated. Changes in the ratio of abundance of these peptides (patent filed Hyperion Biotechnology, San Antonio, Texas) have been shown to be associated with human performance decrement and perception of fatigue during endurance exercise (unpublished data). Briefly, 5 ml of saliva was collected by expectoration into a 15 ml conical tube and stored at −80°C until processed for analysis by liquid chromatography/mass spectroscopy. The saliva sample was taken on the morning of the first day of training before any physical activity was started. Saliva was centrifuged and the supernatant was filtered sequentially through 50,000-molecular weight (MW) and 10,000-MW filters (Microcon; Millipore, Billerica, Massachusetts). The 10,000-MW filtrate was loaded onto a peptide trap column (C8; Michrom, Auburn, California) and eluted in approximately 200 μL of elution buffer. The eluted sample was then dried using a heated vacuum chamber (Centrivap; Labconco, Kansas City, Missouri). The reconstituted sample was injected into the electrospray ionization chamber of an Esquire 3000+ ion-trap mass spectrometer (Bruker, Billerica, Massachusetts). Spectra were visualized, and abundance of peptides was measured using PeakQuest analysis software (Hyperion Biotechnology, San Antonio, Texas). The ratio of abundance, or GGHPPPP:ESPSLIA, is reported as the fatigue biomarker index.

### Statistical Analysis

A classification tree approach was used to develop a predictive model of training success/failure. In some cases, the data were missing and were imputed by generation of values using Monte Carlo simulation on the basis of probability distribution determined by available data. All statistical analysis was done using R 2.9.0 with mice package for the imputation of missing data and the tree package for the classification tree. Briefly, classification trees are created by application of the following steps in the statistical language R 2.9.2 (http://cran.r-project.org/): (1) from a single set of observations (the node), estimate error of a predictive model; (2) search over the set of all possible partitions of the current node, choosing the partitions for which the model's error is reduced as much as possible. If partitioning creates nodes with less than a predetermined number of observations or if the error was reduced by less than a certain amount, the partitioning is stopped; and (3) for each new node created, return to step one and repeat until no more new nodes can be created using limitations set forth in step 2. The model is validated, and sensitivity and specificity were calculated using the leave-one-out method. The method involves taking one observation out of the dataset, fitting the model, and predicting the category the observation left out falls into, either success or failure. The statistical significance of differences in population means of input variables measured in those candidates who succeeded or failed TACP training was made using *t*-tests.

## RESULTS

### Description of the Study Population

Table II shows the reasons for failure in the 4 TACP classes that were evaluated. The total number succeeding is 63 (51.6%) and the number failing is 59 (48.4%). Quitting and medical problems each count for about a third of all failures. About 1 in 6 failures occurs because candidates are unable to meet physical requirements. Administrative and academic reasons make up the remaining failures.

Table III describes the differences between the successful and unsuccessful candidates in terms of age, weight, height, physical performance testing, and fatigue biomarker. Successful candidates are 1.2 years older than the unsuccessful candidates and have better physical performance capability. They are similar with regard to weight, height, physical performance testing, and levels of fatigue biomarker measured at the outset of the training cycle.

Table IV shows the results of the Bar-On EQI testing of TACP candidates. The results show that the following 3 parameters are statistically significantly different: reality test, flexibility, and problem solving. All 3 parameters are lower in successful candidates than in the unsuccessful candidates. Close observation reveals that, with the exception of independence, all test categories and total EQI score are lower in the population of successful candidates compared to those in the population of unsuccessful candidates.

Table V shows the results obtained from the demographic questionnaire. The questionnaire results show that the successful candidates are more educated, spend less time playing video games, are more confident of their ability to complete training, and feel they understand the training requirement better than their unsuccessful counterparts.

**TABLE II.** Reasons Attributed for Failing TACP Training and Prevalence

| Reason | Number | Percent |
|---|---|---|
| Academic: Inability to Meet Standards for Classroom Instruction | 6 | 4.92% |
| Administrative: Disciplinary Problems Leading to Dismissal From Training | 4 | 3.28% |
| Medical: Includes Injuries Sustained During Training Leading to Disqualification | 21 | 17.21% |
| Physical Performance Failure: Candidate Unable to Meet Minimum Physical Performance Requirements, Example, Time to Complete Ruck March | 9 | 7.38% |
| Quit: Candidate Decides Not to Continue | 19 | 15.57% |

The total number of candidates considered is 122. The total number of candidates succeeding is 63 (51.6%) and the number failing is 59 (48.4%).

**TABLE III.** Height, Weight, Age, Physical Fitness, and Fatigue Biomarker Levels in Candidates who Succeed or Fail

| | Success | | | Failure | | | |
|---|---|---|---|---|---|---|---|
| | Median | Mean | SD | Median | Mean | SD | *p*-value |
| Height | 70 | 69.9 | 2.2 | 70 | 70.2 | 2.9 | NS |
| Weight | 173 | 170.8 | 17.6 | 167 | 166.3 | 15.9 | NS |
| Age | 21 | 21.7 | 3.6 | 19.5 | 20.5 | 3.0 | 0.042 |
| Push-ups | 49 | 49.8 | 7.7 | 42 | 43.9 | 8.2 | <0.001 |
| Crunches | 58 | 58.0 | 6.2 | 53.5 | 53.8 | 6.8 | 0.001 |
| Run | 9.9 | 9.9 | 0.7 | 10.4 | 10.1 | 2.2 | 0.006 |
| Pull-ups | 10 | 10.4 | 3.9 | 7.5 | 7.8 | 3.6 | 0.001 |
| Fatigue Biomarker | 0.7 | 12.3 | 33.1 | 0.3 | 15.7 | 40.4 | NS |

**TABLE IV.** Results of the Bar-On EQI

| | Success | | | Failure | | | |
|---|---|---|---|---|---|---|---|
| Trait | Median | Mean | SD | Median | Mean | SD | *p*-Value |
| Total EQI | 97 | 97.4 | 13.9 | 102 | 101.5 | 12.3 | NS |
| Intrapersonal | 100 | 99.4 | 14.4 | 103 | 101.8 | 12.9 | NS |
| Self-regard | 102 | 98.9 | 15.4 | 105 | 102.4 | 12.6 | NS |
| Emotional self-awareness | 98 | 98.6 | 14.1 | 101 | 100.3 | 15.3 | NS |
| Assertiveness | 103 | 103.1 | 14.0 | 103 | 104.9 | 12.5 | NS |
| Independence | 97 | 97.3 | 13.6 | 97 | 96.6 | 12.7 | NS |
| Self-actualization | 101.5 | 100.3 | 14.5 | 106 | 102.8 | 15.1 | NS |
| Interpersonal | 95.5 | 94.7 | 13.1 | 102 | 97.4 | 15.9 | NS |
| Empathy | 89.5 | 91.6 | 13.4 | 98 | 96.0 | 16.3 | NS |
| Social responsibility | 97.5 | 93.3 | 13.7 | 102 | 96.6 | 16.5 | NS |
| Relationships | 99 | 99.1 | 12.8 | 99 | 100.0 | 12.5 | NS |
| Adaptability | 95.5 | 97.3 | 15.5 | 101 | 100.5 | 13.5 | NS |
| Reality testing | 96.5 | 96.9 | 13.2 | 104 | 102.9 | 10.4 | 0.012 |
| Flexibility | 96.5 | 95.9 | 15.0 | 100 | 102.1 | 11.1 | 0.018 |
| Problem solving | 98.5 | 98.5 | 15.4 | 107 | 105.5 | 12.2 | 0.011 |
| Stress Management | 99.5 | 98.1 | 14.2 | 100 | 99.1 | 16.0 | N.S. |
| Stress tolerance | 97.5 | 99.4 | 15.9 | 104 | 103.7 | 13.1 | N.S. |
| Impulse control | 101 | 101.9 | 17.3 | 105 | 106.3 | 13.4 | N.S. |
| General Mood | 96.5 | 98.2 | 13.0 | 97 | 99.5 | 11.1 | N.S. |
| Optimism | 100 | 98.5 | 16.6 | 105 | 102.2 | 12.5 | N.S. |
| Happiness | 100 | 97.5 | 17.8 | 102 | 101.4 | 14.1 | N.S. |

**TABLE V.** Comparison of the Demographic Characteristics of Successful and Failed TACP Candidates

| | Success | | | Failure | | | |
|---|---|---|---|---|---|---|---|
| Question | Median | Mean | SD | Median | Mean | SD | *p*-Value |
| Sports in High School | 2 | 2.2 | 1.4 | 2 | 2.0 | 1.5 | NS |
| How Many Miles Running before 3M | 9 | 9.6 | 7.2 | 9 | 9.9 | 5.9 | NS |
| Hours Strength Training before 3M | 6 | 7.1 | 4.7 | 5 | 7.7 | 6.1 | NS |
| Miles Running before 12 Months | 6 | 8.7 | 9.2 | 8 | 9.2 | 6.1 | NS |
| Hours Strength Training before 12 Months | 6.5 | 7.1 | 4.2 | 6 | 7.8 | 6.1 | NS |
| Body Parts Worked Exercised | 6 | 5.4 | 1.5 | 5 | 5.0 | 1.6 | NS |
| Strength Training Program | 2 | 2.9 | 1.3 | 2 | 2.9 | 1.4 | NS |
| Outdoor Activities | 9 | 8.9 | 4.3 | 8 | 8.9 | 4.7 | NS |
| Living Location | 3 | 2.9 | 1.3 | 3 | 2.8 | 1.1 | NS |
| Position of Leadership | 1 | 0.7 | 0.4 | 1 | 0.7 | 0.4 | NS |
| Time Spent Playing Video Games | 1 | 0.9 | 0.8 | 1 | 1.3 | 0.8 | 0.026 |
| Family in the Military | 1 | 0.8 | 0.4 | 1 | 0.8 | 0.4 | NS |
| Influence of Family Member Serving | 0 | 0.4 | 0.5 | 0 | 0.5 | 0.5 | NS |
| Education | 2 | 2.3 | 1.3 | 1 | 1.8 | 0.9 | 0.018 |
| Confidence in Completing Training | 5 | 4.5 | 0.7 | 4 | 3.6 | 1.2 | <0.001 |
| Understanding of the Training | 3 | 2.9 | 0.8 | 3 | 2.6 | 0.8 | 0.044 |
| Rating of Training as BMT | 3 | 2.8 | 0.8 | 3 | 2.5 | 1.0 | NS |
| Days Lost in Last 12 Months From Injury | 0 | 1.0 | 1.8 | 0 | 1.2 | 1.7 | NS |
| Days Lost in Last 12 Months From Illness | 1 | 1.0 | 1.0 | 2 | 1.5 | 1.4 | NS |

Figure 1 shows the results of comparison of fatigue biomarker values among candidates who succeeded, those who fail for reasons of poor physical performance ability, and those who fail for all other reasons. The results show that, on average, fatigue biomarker levels are 1 order of magnitude less in those who have problems in meeting physical performance requirements than in all others. As the fatigue biomarker index is associated with fatigue produced by physical exhaustion, the finding suggests that individuals performing poorly are to some extent physically exhausted before they begin the training cycle.

### Model Predicting Success and Failure During Training

The classification tree approach determined that only 4 variables, of the 54 that were input initially, were necessary to produce a model that predicts success and outcome in the TACP population (Fig. 2). Inclusion of additional variables did not significantly increase predictive power, whereas dropping 1 of the 4 variables selected had a strong, negative impact on predictive power. The 4 variables included are (1) the runtime for 1.5 miles evaluated at the start of the training cycle, (2) MR or the number of miles run per week during the last year, (3) Ht or the height of the individual summarized, and (4) BM or level of the fatigue biomarker index measured at the outset of training. The order of importance is as follows: run time is more important than MR or running history, which is more important than the fatigue biomarker level and height, which are of approximately equal importance.

Interestingly, of the 4 variables included in the predictive model, only the run time is statistically significantly different, as measured by *t*-test, between the groups that succeed and fail. On the other hand, there are a numerous variables that are statistically significantly different as measured by *t*-test, between successful and failed candidates, that do not contribute in predicting who will succeed and fail using the clustering method. This is not a contradiction because the predictive clustering method groups candidates into several groups, whereas *t*-testing considers only 2, those who fail and those who pass. Ultimately, the clustering method provides better predictions than any single variable or combination judged to be significant by *t*-tests.

The ability of the tree model to correctly identify if a candidate will succeed or fail was tested by the leave-out-one approach described in the methods section. Several parameters were evaluated, including accuracy, sensitivity, and specificity. Accuracy is the percent predicted to pass or fail and that was correctly predicted to pass or fail. Sensitivity is the percent of individuals that were predicted to fail and that actually did fail within a given type or reason of failure. Specificity is the percent predicted succeeding for a given reason divided by the number of those who did not fail for that reason. Table VI shows that the model has excellent specificity, reasonably good accuracy, and to a lesser degree, sensitivity. With regard to medical and physical performance–related causes of failure, the model underestimates the number of candidates falling into these classes. With regard to medical reasons, the model predicts 16 failures when the actual number is 21. And with regard to physical performance–linked failures, the

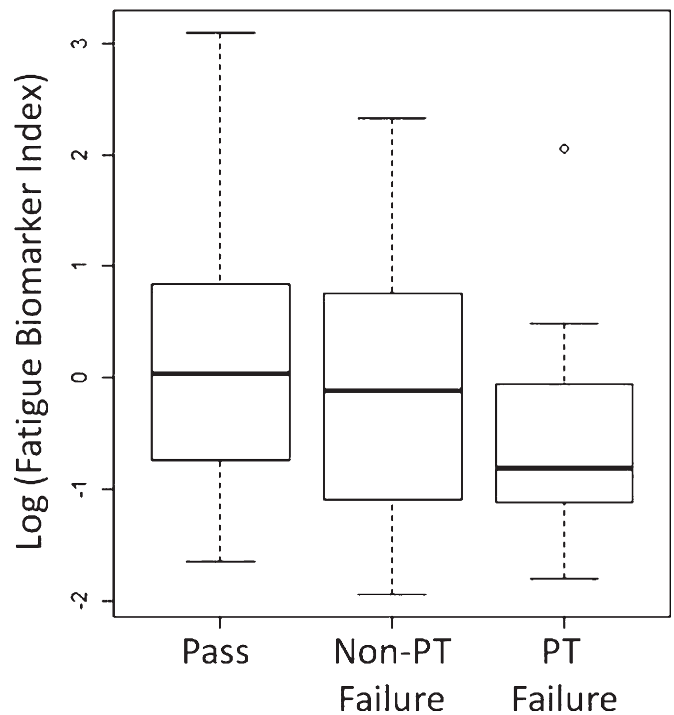


**FIGURE 1.** Biomarker index in candidates who passed and in those who failed.

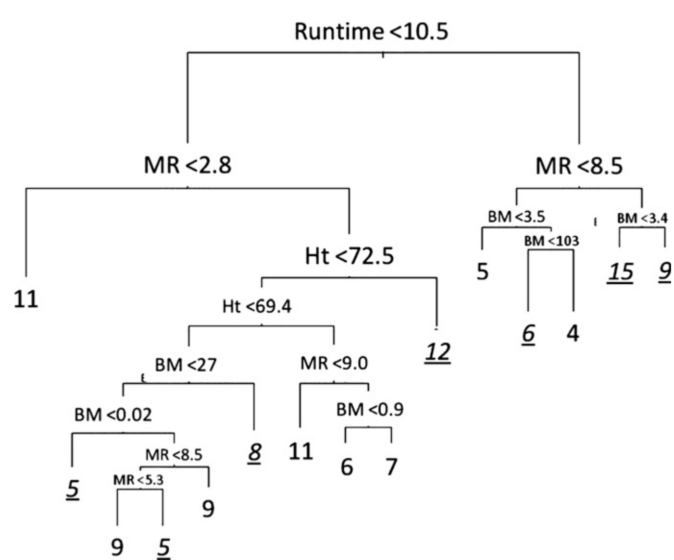


**FIGURE 2.** Graphical depiction of the results of tree analysis. Runtime is the time in minutes required to run 1.5 miles, MR is the self-reported number of miles run per week in the previous 12 months, Ht is the height in inches, and BM is the fatigue biomarker index. Solid numbers represent (eg, **11**) the number of individuals who passed, whereas numbers that are underlined and italicized (eg, *5*) represent the number of failures.

**TABLE VI.** Ability of the Tree Model to Predict Outcome of Training

| Reason | Total Failing (Actual) | Total Failing (Predicted) | Correctly Predicted | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|---|---|
| Academic | 6 | 5 | 3 | 80% | 50% | 98% |
| Administrative | 4 | 5 | 3 | 80% | 75% | 98% |
| Medical | 21 | 16 | 10 | 73% | 48% | 94% |
| Physical Performance Failure | 9 | 5 | 4 | 78% | 44% | 99% |
| Quit | 19 | 21 | 11 | 69% | 58% | 90% |

Accuracy is the percent of individuals who were predicted to pass or fail and who were correctly predicted to pass or fail. Sensitivity is the percentage of individuals who were predicted to fail and who actually did fail within a given type or reason of failure. Specificity is the percent predicted not to fail (succeed) for a given reason divided by the number of those who did not fail for that reason.

model predicts 5 when the actual number is 9. We hypothesize that the reasons for this somewhat poor performance could be attributed to changes in the candidate that occur during the training cycle, which cannot be readily identified with any of the measured parameters at the outset of training. We hypothesize that a frequent cause of this phenomenon may be the cumulative effect of minor injuries, for example, shin splints and blisters that eventually cause the candidate to fail or report for medical disqualification.

## DISCUSSION

Screening applicants for specialized combat training is essential for increasing the training success rate and final deployment of personnel to combat roles. Effective screening methods should identify 3 groups of applicants: those who are likely to succeed in training without intervention, those who are likely to succeed if given suitable support (eg, additional endurance conditioning) before training and those who are unlikely to successfully complete training and should be referred to an alternative career path. In the current study, we have attempted to identify specific applicant characteristics that correlate with success or failure in the Air Force TACP training. Evaluation of demographic data, psychosocial testing, physical performance and activity, and fatigue biomarker index demonstrated that only a limited set of characteristics could be used to predict the success or failure of TACP training. Understanding why some military trainees perform well and others fail has been a topic of interest for military forces around the world. Research in recent years has focused both on psychological and physical performance factors. With regard to basic military training (BMT) failure, Lubin et al[5] showed that negative mood scale as measured on the State Trait-Depression Adjective Check Lists predicts success. In a later article, using a broader test instrument, Lubin et al[6] showed that anxiety, depression, hostility, positive affect, and sensation seeking scales predicted success during Air Force Basic Training. In another study with Air Force BMTs, Carbone et al[7] showed that in 632 candidates a relatively small number of questions from the Minnesota Multiphasic Personality Inventory-2 predicted outcome. In Special Forces, the role that various factors have on training success has been examined previously, though with some emphasis placed once again on psychological testing. In the Norwegian naval special forces, the Rorshach method of evaluating psychological composition and state was shown to accurately predict success.[8,9] In the U.S. Army, 9 elite soldiers were put through training exercises that included both routine and unrehearsed tasks.[10] The findings suggest that the soldiers recover relatively rapidly from physical demands compared to the emotional stress. Collectively, these findings agree broadly with our observation that psychological factors, particularly flexibility, adaptability, and confidence, are associated with training success.

We have shown that there are numerous differences between populations that succeed and fail, but that only a few variables are sufficient to predict training outcome. The successful candidates tend to be older, more educated, more confident, and emotionally resilient. They are also stronger and can run faster. These observations are not altogether surprising. What we found surprising and somewhat counterintuitive is that a very small number of variables are needed to predict success and failure. Of the 4 variables needed to predict outcome, only run time is statistically significantly different between successful and failed groups of candidates. The other variables including distance run per week in the last year before BMT, height, and biomarker are not statistically different. It is important to recognize that selection of these variables was made in an entirely automated fashion and without bias. One explanation for variable selection is that our clustering algorithm identifies distinctly different groups of individuals who fail or pass. Thus, one explanation is that although some variables, for example, time spent playing video games, may be greater in failed individuals, there are no distinct clusters of failed candidates defined by their video gaming behavior.

The salivary biomarker was demonstrated to be a means of identifying candidates at risk of failure. We have demonstrated that reduced levels of the fatigue biomarker index, measured at the outset of the training period, are associated with inability to meet physical performance requirements during training. However, the number of subjects who failed for not being able to meet the requirements is relatively small. Repeated measurements of the fatigue biomarker index made during the training cycle may be useful in identifying candidates who eventually fail. It is possible that some failures for other reasons, including medical disqualification, academic reasons, and quitting, may be due in large measure to fatigue. This proposition would be reasonably straightforward to test by measuring levels in subject over time. More importantly, if the hypothesis that a biomarker approach can be used to identify at-risk individuals is validated, real-time evaluation could be used to provide effective interventions that prevent qualified candidates from failing. The idea of intervening selectively is controversial, but has been explored in the military training context.[11]

## ACKNOWLEDGMENTS

We acknowledge the cooperation of AFSOC 720th Special Tactics Group and the TACP candidates who participated in this study. This work was supported by the USAF through contract no. FA3047-08-D-0018 awarded to Hyperion Biotechnology, Inc.

## REFERENCES

1. Wilson JR: Not too close: putting ordnance where ground forces want it. Armed Forces J 2004; 141(1): 42–4.
2. DePalo, Lee K: USAF Combat Search and Rescue: Untapped Combat Power. Maxwell AFB, AL, Air University Press, 2005.
3. Theisen, Eric E: Ground-aided Precision Strike: Heavy Bomber Activity in Operation Enduring Freedom. Maxwell AFB, AL, Air University Press, 2003.
4. Bar-On R: The Bar-On model of emotional–social intelligence (ESI). Psicothema 2006; 18(Suppl): 13–25.
5. Lubin B, Fiedler ER, Van Whitlock R: Mood as a predictor of discharge from Air Force basic training. J Clin Psychol 1996; 52(2): 145–51.
6. Lubin B, Fiedler ER, Van Whitlock R: Predicting discharge from Air Force basic training by pattern of affect. J Clin Psychol 1999; 55(1): 71–8.
7. Carbone EG, Cigrang JA, Todd SL, Fiedler ER: Predicting outcome of military basic training for individuals referred for psychological evaluation. J Pers Assess 1999; 72(2): 256–65.
8. Hartmann E, Sunde T, Kristensen W, Martinussen M: Psychological measures as predictors of military training performance. J Pers Assess 2003; 80(1): 87–98.
9. Hartmann E, Grønnerød C: Rorschach variables and Big Five scales as predictors of military training completion: a replication study of the selection of candidates to the naval special forces in Norway. J Pers Assess 2009; 91(3): 254–64.
10. Thomas JL, Adler AB, Wittels P, Enne R, Johannes B: Comparing elite soldiers' perceptions of psychological and physical demands during military training. Mil Med 2004; 169(7): 526–30.
11. Williams A, Hagerty BM, Andrei AC, Yousha SM, Hirth RA, Hoyle KS: STARS: strategies to assist navy recruits' success. Mil Med 2007; 172(9): 942–9.