

**Comparison of Measured and Indicated Calories Among New and Used Life Fitness CLSC  
Upright Cycles**

**James A. Hodgdon, Ph.D.  
Eagle Applied Sciences  
San Antonio, TX 78248**

**Sofie Blicher, M.S.  
San Diego State University  
San Diego, CA 92182**

**Michael J. Buono, Ph.D.  
San Diego State University  
San Diego, CA 92182**

**Jay Heaney, MA  
Warfighter Performance Department  
Naval Health Research Center  
San Diego, CA 92152**

## **Disclaimer Statements**

### *Authors disclaimer*

The views expressed in this article reflect the results of research conducted by the author and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the United States Government.

### *Source of Support*

This work was supported/funded by work unit number 60601.

### *IRB statement*

The study protocol was approved by the San Diego State University (SDSU) Institutional Review Board in compliance with all applicable federal regulations governing the protection of human subjects. NHRC's Human Research Protection Office provided a HRPO review as this research was conducted by SDSU.

### *Copyright Statement*

I am a federal employee of the United States government. This work was prepared as part of my official duties. Title 17 U.S.C. 105 provides that copyright protection under this title is not available for any work of the United States Government.' Title 17 U.S.C. 101 defines a U.S. Government work as work prepared by an employee of the U.S. Government as part of that person's official duties.

## Summary

### *Background*

In 2007, the Navy Physical Readiness Program Office (N-135F) authorized the use of stationary cycle (bike) test as an alternative to the 1.5-mile run in the Navy Physical Fitness Assessment (PFA). The number of calories expended in 12 minutes is used with body weight to predict equivalent time for a 1.5-mile run.

Previous work with elliptical trainers showed that although calorie indicators vary in accuracy among bike models, a comparison of 5 models of a particular elliptical trainer showed identical cumulative calories indicated for the same time, speed, and resistance. Comparable data have not been obtained for stationary cycles. To that end, the Exercise and Nutritional Sciences Department (ENS) of San Diego State University (SDSU) was asked by NHRC to conduct validity trials of new and used Life Fitness CLSC upright cycle ergometers. Here, we report the results of those trials.

### *Methods*

Life Fitness Corporation (Rosemont, IL) provided 3 new model CLSC stationary cycles for testing (labeled bikes #1 - #3). An additional 3 used CLSC stationary cycles were obtained from a Navy Fitness Center in the San Diego area (bikes #4 - #6). Ten participants were recruited from the ENS students and faculty (5 men and 5 women). Gender groups differed in stature and weight, but not in mean age. Each participant worked at three different work rates on each of the stationary cycles. Each work bout lasted approximately 12 minutes. For the last 6 minutes of each bout, calories expended were determined from open-circuit spirometry measures (**kcal<sub>spiro</sub>**). For the same period, the number of expended calories indicated on the cycle display was also recorded for comparison (**kcal<sub>indic</sub>**).

The regression of **kcal<sub>spiro</sub>** on **kcal<sub>indic</sub>** was determined for each bike. The regression for each bike was inspected to determine whether or not the 95% confidence interval of the slope included the value 1. This is the Navy performance requirement to qualify a bike to be used for Physical Readiness Assessment.

Individual bike regressions were then compared in pairwise fashion between bikes. To carry out these comparisons, we used method of Rogosa (1980, 1981). This analysis is based on the difference function of the two regressions and allows an evaluation of equality of slope and calorie intervals over which the bikes are the same or differ in their indicated calorie response.

### *Results*

The mean values for **kcal<sub>spiro</sub>** or **kcal<sub>indic</sub>** did not differ among bikes ( $F_{5,173} = 0.093$ ,  $P = 0.993$ ; and  $F_{5,173} = 0.330$ ,  $P = 0.894$ , respectively), but the indicated calories were greater than the measured calories for each bike (mean difference = -5.48 kcal,  $P < 0.033$ ).

Univariate ANOVA of **kcal<sub>spiro</sub>** with bike number as the grouping variable and **kcal<sub>indic</sub>** as the covariate revealed a significant effect of bike number ( $F_{5,172} = 5.93$ ,  $P < 0.001$ ). Pairwise comparisons of the estimated marginal means found the mean for bike #5 to differ from those of all the other bikes (mean  $P = .003$ ), while significant differences were not found for comparisons of the means among any of the other bikes.

As expected, the regressions among the new bikes did not differ significantly from one another. There was also agreement of regressions of new bikes #1, #2, and #3 with used bike #4. Additionally, new bike #1 and #2 aligned with used bike #6. Mimicking the ANOVA results, the regression of bike #5 differed significantly from that of any other bike. The regressions for the older bikes #4 and #6 had regressions that were indistinguishable from new bike #2 or #3, respectively. There was no general pattern for change in the regression slopes and/or intercept between the bike groups. That older bike #5 differs from all the others suggests that there can be meaningful changes in performance with use. A program of periodic validation should be implemented to identify bikes that vary from “like-new” performance.

As mentioned above, the current Navy performance standard for stationary bikes is that the 95% confidence interval for the slope of the regression of **kcal<sub>spiro</sub>** on **kcal<sub>indic</sub>** includes the value 1. Three of the bikes tested failed to meet this criterion, #3 in the new group and #4 and #5 in the old group. The regressions all have similar standard errors for the slope and for the regressions as a whole. If the value of the slope was 0.907 or less, the 95% confidence interval did not include 1. That one of the bikes meeting this condition was in the new group implies new bikes of this model include bikes that do not meet the Navy specifications. It also implies that testing a single bike as an accurate representative of the model is not satisfactory.

Perhaps future testing should require several examples of a particular model to be tested in order to assess inter-bike variation. If examples are found that do not meet Navy standards, then the model should be rejected until its deficiencies can be overcome.

## **Background**

In 2007, the Navy Physical Readiness Program Office (N-135F) authorized the use of a stationary bike test as an alternative to the 1.5-mile run in the Navy Physical Fitness Assessment [PFA; Chief of Naval Operations (2007)]. A test subject would pedal on a stationary bike for 12 minutes. The number of calories generated during the 12 minutes are recorded and used with body weight to predict a 1.5-mile run time (Hodgdon et al., 2007).

When new exercise ergometers are considered for use in the Physical Readiness Test (PRT), validation of caloric expenditure accuracy is required. Only validated and approved PRT ergometers are entered into the Physical Readiness Information Management System (PRIMS). PRIMS is governed by the directive issued by the Chief of Naval Operations (OPNAVINST 6110.1 series) and provides a standardized Navy database to monitor and track the progress of Navy active duty and reserve personnel Physical Fitness Assessment (PFA) data. Since 2002, PRIMS is the sole Navy authoritative data source that houses PFA data for command information, authorization information, and member personnel data including PRT data.

Previous work at the Naval Health Research Center (NHRC) has shown that identical data can be obtained from five cardiovascular exercise (cardio) devices of the same model (elliptical trainers) (Life Fitness CT 9500HR) (Parker, Griswold & Vickers, 2006). Results showed identical caloric expenditure for same speed and resistance settings across all 5 devices. This initial cardio device study established the validation procedure of recording actual calorie expenditure via measurement of oxygen consumption and comparing the cumulative number of measured calories with the cumulative number of observed calories displayed from the cardio device. However result from a follow-on evaluation of three different brands and models of cardio devices were shown to differ in their accuracy of calorie indicators (Vickers, Griswold & Hodgdon, 2006). Subsequent evaluations have indicated that the accuracy of calories generated from cardio devices can change with use and over time (Buono, Heaney, Hodgdon, 2014, unpublished data). Due to the uncertainty of cardio devices providing a consistently accurate generation of calories with long term use and manufacturing age, an evaluation of multiple new and used models of the same device was proposed.

Over the last several years the Life Fitness CLSC stationary bike, an approved Navy alternative cardio device, has been the most popular bike used with PRT assessments. However an

evaluation of multiple identical models of new and used Life Fitness CLSC upright bike ergometers has never been investigated. The purposes of this validation study were twofold: (1) to determine accuracy and validity of several new bike ergometers, and (2) to determine if there was disparity in caloric expenditure measurements between new unused and older used bike ergometers.

## Methods

Life Fitness Corporation (Rosemont, IL) provided 3 new model CLSC stationary bikes for testing (bikes #1 - #3). An additional 3 used CLSC stationary bikes were obtained from a Navy Fitness Center in the San Diego area (bikes #4 - #6).

## Subjects

Ten participants (5 males, 5 females) were recruited from a local university. Their physical characteristics are provided in Table 1. There were sex-related differences in height and weight, but not in mean age.

**Table 1, Physical Characteristics<sup>1</sup>**

	Males ( <i>n</i> =5)	Females ( <i>n</i> =5)	Total
Age	26.2 (2.3)	25.8 (1.5)	26.0 (1.8)
Height (in)	71.25 (3.87)	64.4 <sup>2</sup> (1.77)	67.83 (4.59)
Weight (lb)	170.9 (18.9)	124.8 <sup>2</sup> (6.0)	147.9 (1.8)

<sup>1</sup>Values shown are means (1SD), <sup>2</sup>Differs from male mean ( $P<0.05$ )

## Procedures

Participants reported to a university laboratory on 6 occasions for testing. During their first visit, they were informed about the nature of the study, the procedures and risks involved, and given the opportunity to volunteer. Tests across the six cycles were randomized and limited to one bike trial per day. During each test, the subject worked on the bike at three different power settings and the work intensity increased with each successive work bout. Heart rate (HR) was monitored to ensure that the first workload was between 95-105 bpm, the second workload HR was between 110-120 bpm, and the third workload was between 125-135 bpm. Each bout was approximately 12 minutes in length. A 5-minute break was provided between bouts, during

which the participant was allowed to remove the spirometry mask and nose clip. During the work bout, expired gas was collected and analyzed using a Parvo™ True One open-circuit spirometry system (Parvo Medics, Sandy, UT) to provide measurement of the rate of oxygen consumption ( $\dot{V}_{O_2}$ ). The system also calculated the calories expended from the oxygen consumption values using the Weir (1949) formula. The total calories expended during minutes 6 to 12 of the test from open-circuit spirometry were recorded and doubled to provide a comparable 12-minute calorie count displayed on the bike.

### Analysis

The focus of this project was on the accuracy of the bike generated calories provided on the display screen with the actual calories measured from oxygen consumption (difference between measured calories and displayed calories). Additionally, the validity of the displayed calories was evaluated to ensure that a change in measured calories would result in a change in bike generated calories. Therefore, a regression analysis of calories from oxygen consumption on time was computed for each test. The 95% confidence interval for the slope was checked to ensure that it included the value zero. This was done to minimize the possibility that the participant was working above his/her ventilatory threshold. Above ventilatory threshold, the rate of oxygen consumption would increase without a change in physical work rate, and provide an unstable measurement of the metabolic cost of the work. One test was excluded because the regression slope did not meet this criterion.

The number of calories displayed on the bike at minutes 6 and 12 of the work bout were also recorded, and the difference calculated to provide the 6-minute bike estimate of calories expended during this period. The total expended calories from spirometry (**kcal<sub>spiro</sub>**) from minute 6 to minute 12 were used as the criterion against which expended calories indicated on the bike (**kcal<sub>indic</sub>**) were compared. The regression of **kcal<sub>spiro</sub>** on **kcal<sub>indic</sub>** was determined for each bike. These regressions were compared in pairwise fashion between bikes. To carry out these comparisons, we used the method of Rogosa (1980, 1981).

The typical analysis of covariance approach to evaluate the parallelism of regression lines requires the assumption of parallel slopes to test differences in the estimated marginal means of the regressions. The Rogosa approach avoids this assumption. This is important because rarely are the slopes of regression lines that are being compared exactly equal. In analysis of

covariance, a difference between the intercepts of the lines is calculated and assumed to apply for all values of the independent variable. The more common reality is that the slopes of the regressions differ and that the offset between regressions varies as a function of the independent variable.

Rogosa's method is based on analysis of the difference function between two regressions. The difference function =  $(b_{0_1} - b_{0_2}) + (b_{1_1} - b_{1_2}) \times \mathbf{X}$ , where  $b_{0_1}$  and  $b_{0_2}$  are the intercepts for regressions 1 and 2,  $b_{1_1}$  and  $b_{1_2}$  are the regression slopes, and  $\mathbf{X}$  is the independent variable. The confidence intervals of the difference function are used to identify regions of significant difference between the regression lines being compared. Values of the difference function that differ significantly from zero are called Johnson-Neyman (J-N) regions of significance after the developers of this concept (Johnson & Neyman, 1936). Such regions are defined by the intersection of the confidence intervals of the difference function with zero.

The nature of the J-N region of significance depends on the coefficients of the difference function. If the slope of the difference function differs greatly from zero, the J-N regions indicate values of the independent variable beyond which the regressions are not similar. This implies an "interior" region over which the regressions do not differ significantly. If the slope of the difference function is near zero, the J-N region describes an interior set of values over which the regressions differ, but beyond which significant difference cannot be demonstrated.

Comparison of regressions was carried out sequentially for pairs of elliptical trainers. For convenience, the difference between regression pairs was evaluated at the grand mean of the prediction from both regressions. In this way, we get an indication of the offset magnitude between regressions near the center of the range covered by regressions.

To avoid capitalization on chance by making multiple comparisons (in this case, 15), the Bonferroni correction (Abdi, 2007) was applied. The critical P value was set at  $0.05/15 = 0.003$  for the comparisons between regressions.

Means and standard deviation and regression calculations were carried out using PASW Statistics, ver. 18 (SPSS Inc., Haverhill, MA). Comparison of regressions was carried out using a Microsoft Excel™ spreadsheet (Microsoft Corporation, Redmond, WA) using the calculations presented by D'Alanzo (2004).

## Results

Table 2 shows the mean, minimum, and maximum measured and indicated calories for each bike, as well as the difference between **kcal<sub>spiro</sub>** and **kcal<sub>indic</sub>**. Neither the mean values for **kcal<sub>spiro</sub>** nor **kcal<sub>indic</sub>** differed among bikes ( $F_{5,173} = 0.093$ ,  $P = 0.993$ ; and  $F_{5,173} = 0.330$ ,  $P = 0.894$ , respectively), but the indicated calories were greater than the measured calories for each bike ( $P < 0.033$ ).

	Bike	Measured Calories	Indicated Calories	Difference <sup>2</sup>
New	#1	34.09 (17.83, 72.66)	39.79 (24, 81)	-5.78
	#2	33.52 (18.05, 79.74)	39.00 (24, 85)	-5.48
	#3	35.10 (19.42, 72.37)	40.10 (24, 83)	-5.00
Used	#4	35.10 (19.84, 82.45)	40.87 (25, 93)	-5.88
	#5	35.33 (17.28, 78.05)	43.57 (24, 93)	-8.24
	#6	35.51 (19.42, 87.28)	40.83 (24, 91)	-5.33

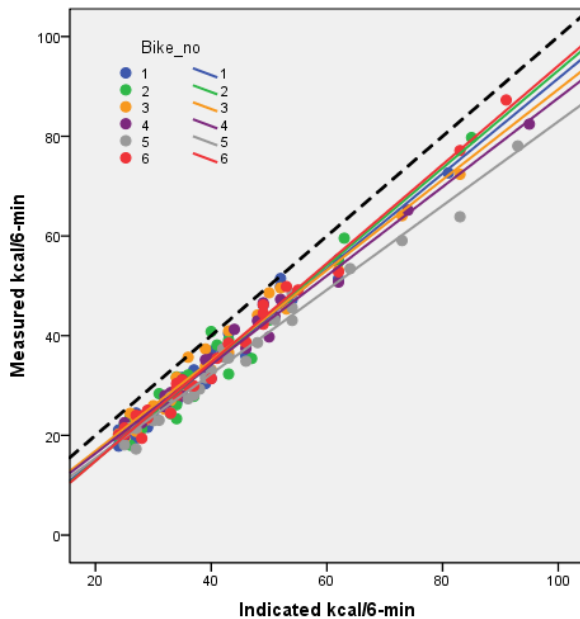
<sup>1</sup>Values shown are means with (minimum, maximum)  
<sup>2</sup>Measured – Indicated Calories

Univariate ANOVA of **kcal<sub>spiro</sub>** with bike number as the grouping variable and **kcal<sub>indic</sub>** as the covariate revealed a significant effect of bike number ( $F_{5,172} = 5.93$ ,  $P < 0.001$ ). Pairwise comparisons of the estimated marginal means found the mean for bike #5 to differ from those of all the other bikes (mean  $P = .003$ ), while significant differences were not found for comparisons of the means among any of the other bikes.

Table 3 provides the coefficients for the regression of **kcal<sub>spiro</sub>** on **kcal<sub>indic</sub>** for each bike. Also included are the squared correlation coefficient values and the standard error of estimate (SEE).

	Bike	Slope (kcal·kcal <sup>-1</sup> )	Intercept (kcal)	r <sup>2</sup>	SEE
New	#1	0.956 (0.882, 1.029)	-3.938 (-7.019, -0.859)	0.963	2.55
	#2	0.978 (0.898, 1.058)	-4.635 (-7.938, -1.332)	0.957	2.87

	#3	0.907 (0.847, 0.967)	-1.265 (-3.809, 1.279)	0.972	2.21
Used	#4	0.891 (0.845, 0.937)	-1.431 (-3.447, 0.585)	0.982	1.87
	#5	0.844 (0.797, 0.891)	-1.443 (-3.627, 0.740)	0.980	2.09
	#6	0.991 (0.948, 1.035)	-4.960 (-6.861, -3.059)	0.987	1.82
<sup>1</sup> Values shown are coefficients (95% confidence limits)					



**Figure 1. Relationship between indicated and measured calories over a 6-min interval for each of the bikes used in this study. Dotted line is the line of identity for the comparison.**

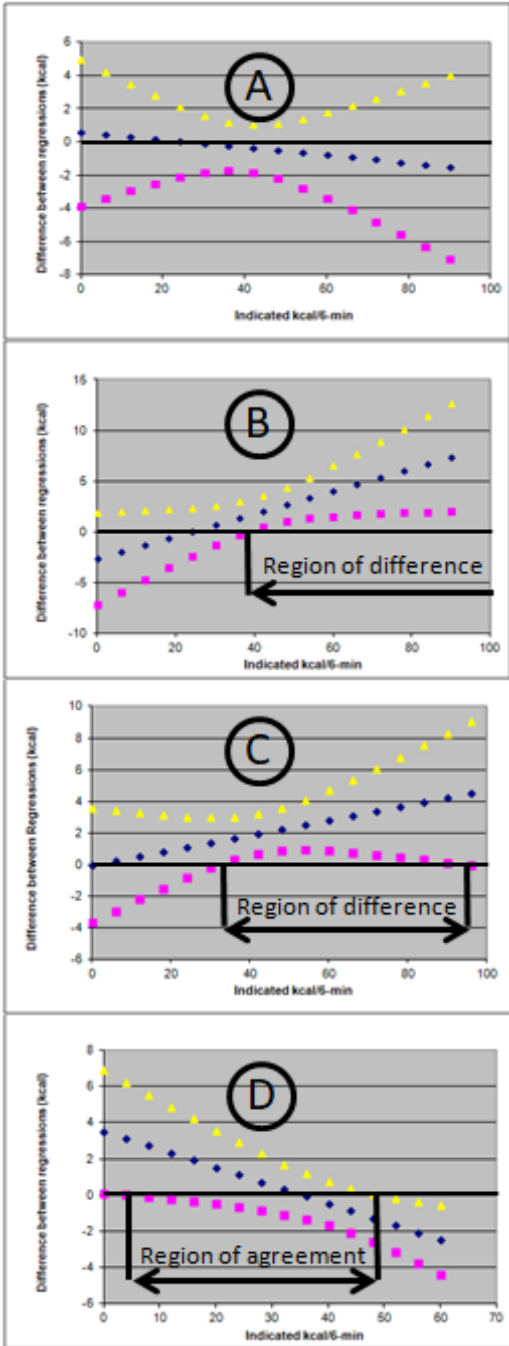
Figure 1 shows the individual relationships between indicated and measured calories for each of the bikes used in this study. The solid lines indicate the least squares linear regression of measured on indicated calories. The dotted line represents the line of identity. As noted above the indicated calorie values are greater than the measured values for all the bikes.

Table 4 provides the results of the comparison of regressions for each pair of bikes. For each comparison, the table lists a comparison of the slopes of the

regressions as well as the presence or absence of a J-N region of non-agreement and the boundaries of that region, when present. If the slopes are equal and a J-N region is not present, the regressions do not differ significantly. The table has been color coded to make it easier to read. Regressions that do not differ significantly are shaded in green. Regressions that have a J-N region of inequality are indicated in red. Regressions that differ from the general pattern of unequal slope with the presence of a region of inequality are shaded in orange.

**Table 4. Comparison of Regressions**

	Comparison (Bike #)	Regression Slopes =?	Johnson- Neyman region exists?	Lower Boundary	Upper Boundary	Interpretation
New	#1 v. #2	yes	no			
	#1 v. #3	yes	no			
	#2 v. #3	yes	no			
Used	#4 v. #5	yes	yes	32.1	94.7	regressions <u>differ</u> between boundaries
	#4 v. #6	no	yes	4.7	48.8	regressions <u>agree</u> between boundaries
	#5 v. #6	no	yes	-0.5	33.7	regressions <u>differ</u> above 33.7 kcal
Btwn	#1 v. #4	yes	no			
	#1 v. #5	no	yes	-253.1	37.3	regressions <u>differ</u> above 37.3 kcal
	#1 v. #6	yes	no			
	#2 v. #4	yes	no			
	#2 v. #5	no	yes	-60.2	36.6	regressions <u>differ</u> above 36.6 kcal
	#2 v. #6	yes	no			
	#3 v. #4	yes	no			
	#3 v. #5	yes	yes	27.6	126.3	regressions <u>differ</u> between boundaries
#3 v. #6	no	no				



**Figure 2.** Graphs of difference function with confidence intervals for different comparison patterns. Panel A, equal slopes, no Johnson-Neyman region. Panel B, unequal slopes, Johnson-Neyman limit above which regressions differ. Panel C, equal slopes, Johnson-Neyman region within which the regressions differ. Panel D, unequal slopes, Johnson-Neyman region within which the regressions do not differ.

A first observation from the Table 4 results is that the new stationary bikes all respond in the same way. None of the comparisons among the new bikes yield statistically different regressions. Secondly, we see that all of the used bikes (#4, #5, and #6) show  $\text{kcal}_{\text{spiro}}$  for  $\text{kcal}_{\text{indic}}$  relationships that differ with those of the other used bikes. Thirdly, inspection of the results indicates that the bike #5 regression differs significantly from those of all other bikes.

## Discussion and Conclusions

### Effects of Bike Use.

There are 3 different patterns of non-congruence in these data. These patterns are shown in Figure 2. Each panel shows the difference function (blue symbols) for one of the comparisons with its 95% confidence intervals (shown in yellow symbols for the upper limit and pink symbols for the lower limit). J-N regions are determined by the interaction of the confidence limits and the zero value on the y-axis. Panel A of Figure 2 shows the difference function for a comparison in which the slopes are equal and there are no J-N regions (bike #1 with bike #2). The slope of the difference function is modest,  $-0.023 \text{ kcal}_{\text{indic}} / \text{kcal}_{\text{spiro}}$  and the confidence intervals do not cross the  $Y=0$  axis value. Panel B shows the pattern where the regressions diverge above the greater J-N limit. In this case (#1 with #5) the lower J-N limit is negative and has no meaning in this measurement system. The lower

confidence interval of the function crosses the  $Y=0$  line at 37.3  $\text{kcal}_{\text{indic}}$  and remains greater than zero as  $\text{kcal}_{\text{indic}}$  increases. There are 3 examples of this pattern among the comparisons (see Table 4). Panel C depicts the pattern in which there is a J-N region within the function for which the regressions differ. There are two examples of this pattern among the comparisons. In each case, the lower confidence limit exceeds zero and then as the confidence interval increases with increasing indicated calories the lower limit again crosses  $Y=0$ . The interval between the zero crossings defines a region for which the regressions differ. Because this region is limited due to increasing uncertainty associated with the difference function, the practical application is the same as for patterns shown in Panel B. That is, the regressions diverge for values greater than the J-N limit. Panel D shows a pattern wherein the J-N region is one of agreement between the two regressions. In this instance (#4 with #6), the lower confidence limit crosses  $Y=0$  to define the beginning of the region. The function has a negative slope and eventually the upper confidence limit crosses  $Y=0$ . For the region between the two crossings the confidence interval includes the value  $Y=0$ , and is therefore a region of agreement between the two regressions. The practical application here, too, is that above a certain value the regressions differ. There was also one unique result, that of #3 with #6. In this case there was no J-N region defined, but the slopes of the regressions differed significantly. The difference function resembles that in Panel A with no crossings of the confidence intervals and  $Y=0$ , but with the slopes of the regressions differing significantly.

As should be the case, the regressions among the new bikes did not differ significantly from one another. There was also agreement of regressions of new bikes #1, #2, and #3 with used bike #4, and #1 and #2 with used bike #6. Bikes #4 and #6 each had regressions similar to examples in the new bike group. The one clear finding was that used bike #5 had a different calorie reporting response than any of the other bikes.

There was no general pattern for change in the regression slopes and/or intercept between the bike groups. That bike #5 differs from all the others suggests that there can be meaningful changes in performance with use. A program of periodic validation should be implemented to identify bikes that vary from “like-new” performance.

### **Relationship to Current Bike Acceptance Standards.**

Current Navy standards for acceptance of a stationary bike to be used in the Physical Readiness Assessment (PFA) are based on the results of tests comparing calories measured during exercise with those indicated on the bike (Navy Electronic Commerce Online, 2010). The best fit line is calculated to predict the measured calorie uptake from the calorie intake indicated on the stationary bike. For the stationary bike to be acceptable, the slope of this line (the relative change in the measured calories for each unit of change in the indicated calories) must equal to 1. This means that for every change in the value of the indicated calories, there is an equal change in the value of the measured calories. Values greater than 1 imply that measured calories expended during a work bout are greater than the indicated calories for that bout. Conversely, a value less than 1 implies that measured calories expended are less than those indicated on the bike. Of course, the measured value of the slope is seldom equal to 1 exactly. A decision needs to be made whether or not this slope is “close enough” to the value 1 to treat it as equal to 1. The decision is based on the statistical confidence limits of the slope. (For a discussion of the confidence interval of the regression slope, see any basic statistics text.) Navy policy has been that when the 95% confidence limits of the slope (the range of values for the slope for which there is only a 5% probability that the “true” value of the slope lies outside the range) includes the value 1, it is deemed close enough to estimate it as actually being 1.

When this is the case, the mean difference between actual calories expended and calories indicated on the bike is calculated. This difference is called an offset and the reported calories can be corrected to match actual calories by adding this offset. The Navy sets no limits on the size of the offset when determining the acceptability of a particular stationary bike. The only requirement is that the confidence interval for the slope of the measured on indicated calories include the value 1.

Current Navy standards also require the regression of measured on indicated calories to be based on at least 75 measurements. In previous testing of stationary bikes, this has been achieved by having 25 participants ride each bike at three different workloads to yield 75 total tests.

Our findings in these tests point out a limitation of the Navy’s testing approach. Currently, the Navy tests only a single bike to represent a particular model. If these bikes were being tested for Navy acceptance, only two of the three new bikes would have been found acceptable. The 95%

confidence interval of the regression slope included 1 for bikes #1 and #2, but not bike #3. If bike #3 was tested as the sole representative of its class, the class would have been rejected. If bikes #1 or #2 had been tested, the noncompliance of bike #3 would not have been detected.

Because we found equal slopes in the paired regression comparisons among the new bikes, we tried pooling the  $kcal_{\text{spiro}}$  and  $kcal_{\text{indic}}$  values for all three new bikes to determine if the pooled regression would meet the Navy standard. It does not. The pooled regression slope was 0.949 with a confidence interval of 0.907 to 0.986. We also tried increasing the confidence interval to be the 99% confidence interval (the range of values for the slope for which there is only a 1% probability that the “true” value of the slope lies outside the range). The confidence interval for bike #3 under these conditions was 0.826 to 0.988, still not qualifying.

Bike No.	Slope of the Regression	Standard Error of the slope	Slope Include 1?
1	0.956	0.038	yes
2	0.978	0.041	yes
3	0.907	0.031	no
4	0.891	0.023	no
5	0.844	0.024	no
6	0.991	0.022	yes

Table 5 shows results for the regression slope, the standard error of that slope, and if the 95% confidence interval of the slope includes 1. The only pattern that appears is that when the slope is equal to .907 or less, the confidence interval does not include 1. Further, it appears that 2 of the 3 new bikes produced by this manufacturer yielded results within this range, while 2 of the 3 used bikes did not.

If the new bikes represent the mechanical tolerances of the manufacturer’s production process, then finding that a single bike meets specifications does not assure that all bikes of the same model do. Perhaps future testing should require several examples of a particular model to be tested in order to assess inter-bike variation. If examples are found that do not meet Navy

standards and they differ so greatly that techniques such as pooling the regressions across the sample are not sufficient to show group compliance, then the model should be rejected until its deficiencies can be overcome.

In summary, the comparisons of new and old models of this stationary bike do not show a consistent pattern of change with use. In fact two of the three “old” bikes performed like their new counterparts. There was however one old bike (#5) that differed in performance from the other 5. This suggests a need for periodic testing of stationary bikes to determine that they are still performing within specs.

## References

- Abdi H (2007). The Bonferroni and Šidák corrections for multiple comparisons. in Salkind N.J. (ed.). *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage.
- Chief of Naval Operations (2007). *Naval Administrative Message NAVADMIN 011/07*, 22 January.
- D’Alonzo KT (2004). The Johnson-Neyman Procedure as an Alternative to ANCOVA. *West J Nurs Res*. 26(7): 804–812.
- Hodgdon JA, Hervig L, Griswold L, Terry J, Le C, Sausen K, Miller P (2007). 12-minute stationary bike performance as a predictor of 1.5-mile run time. *Naval Health Research Center Technical Report No. 07-41*, San Diego, CA.
- Parker SB, Griswold L, Vickers, RR, Jr (2006). Development of an Elliptical Trainer Physical Fitness Test. *Naval Health Research Center Technical Report No. 06-06*, San Diego, CA.
- Vickers RR, Jr, Parker SB, Hodgdon JA (2006) A Comparison of Three Models of Elliptical Trainer. *Naval Health Research Center Technical Report No. 06-31*, San Diego, CA.
- Johnson PO, Neyman J (1936). Tests of certain linear hypotheses and their application to some educational problems. *Stat Res Memoirs* 1: 57-93.
- Navy Electronic Commerce Online, (2010)  
<https://www.neco.navy.mil/synopsis/detail.aspx?id=289358>. Accessed 08 April 2012.
- Rogosa D (1980). Comparing nonparallel regression lines. *Psych Bull* 88(2): 307-321.

Rogosa D (1981). On the relationship between the Johnson-Neyman region of significance and statistical tests of parallel within-group regressions. *Ed Psych Meas* 41: 73-84.

**REPORT DOCUMENTATION PAGE**

*Form Approved  
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD-MM-YYYY)</b>		<b>2. REPORT TYPE</b>		<b>3. DATES COVERED (From - To)</b>	
<b>4. TITLE AND SUBTITLE</b>				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b>	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b>				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b>					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b>					
<b>15. SUBJECT TERMS</b>					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>
<b>a. REPORT</b>	<b>b. ABSTRACT</b>	<b>c. THIS PAGE</b>			<b>19b. TELEPHONE NUMBER (Include area code)</b>