

# A Regression Solution to the Problem of Criterion Score Comparability

**William M. Pugh**  
**Naval Health Research Center**

When the criterion measure in a study is the accumulation of responses or behaviors for an individual over time, there may be a problem in obtaining comparable scores if all individuals cannot be observed for the same amount of time. Past solutions to this problem have included prorating the criterion score and deleting individuals lacking a portion of the criterion information. In the present paper the problem of criterion score comparability is viewed as a missing data problem, and a solution based upon linear regression is proposed. The re-

gression solution was evaluated by comparing it to the prorating and deletion methods in a monte carlo analysis. The population data were an actual data set collected during a study of illness aboard U.S. Navy ships, and a series of samples were created by randomly deleting individuals' records and portions of records. Results of the analyses suggested that the regression method is more efficient than the alternative methods. Application of the regression method and the results obtained to other research settings is discussed.

Frequently, the dependent or criterion variable in an investigation is collected over an extended period of time. For example, in a study of job performance the number of production errors each worker made during a month might be counted. Or, in an epidemiological investigation, the number of illness episodes for an individual may be accumulated over a year or even five years. In such situations it is likely that individuals will enter and leave the study population during the course of the investigation, and when this occurs, scores for individuals are not comparable because they reflect different periods of exposure as well as the propensity to make errors or to become ill.

There are two common remedies taken to overcome this problem with criterion score comparability. One is to control for the length of time people were observed by eliminating any individual not present during the entire study from the data analysis. The other remedy is to divide each criterion count by the length of the observation period, the result often being expressed in terms of a rate (e.g., illness rate or error rate).

The purpose of the present paper is to show that the problem of criterion score comparability can be viewed as a missing data problem in a multivariate analysis. This perspective was gained by considering the total observation period as a continuum that can be divided into several intervals. Viewed in this way, each person's criterion data consist of the set of criterion tallies corresponding to the time

intervals. Taking this position is useful because it provides a basis for evaluating the two solutions discussed above; and from the vantage point obtained, the drawbacks of both procedures become apparent. On the one hand, deleting all the data obtained for a person because of one or two missing data points can result in the waste of large amounts of information (Frane, 1976; Gleason & Staelin, 1975) and, as argued by Galdi and Bonato (1977), that exclusion can introduce bias into the criterion. On the other hand, the practice of computing rates or prorating data appears flawed when viewed from the missing data perspective. Functionally, the prorating procedure treats data from complete observations as a perfect predictor of the unobserved data. Although past and future observations may be related, rarely would a perfect relationship be expected.

In addition to providing a basis for evaluating the deletion and prorating procedures, viewing the problem of criterion comparability as a missing data problem sets the stage for applying procedures designed to handle the problem of incomplete data in a multivariate analysis. Generally, the techniques that have been developed (Buck, 1960; Edgett, 1956; Frane, 1976; Glasser, 1964; Gleason & Staelin, 1975; Rubin 1977; Timm, 1970) use the available information to compensate for the incomplete data by taking advantage of the redundancy among measures. In two studies (Gleason & Staelin, 1975; Timm, 1970) designed to compare the performance of different techniques for handling missing data, it was found that regression techniques, such as the one proposed by Buck (1960), were effective. It was noted, however, that the regression method in which a linear equation is used to predict missing data points from the available data tended to be slow and computationally difficult. Before applying one of these methods to the present problems, it should be pointed out that, for the most part, the existing techniques were designed to handle the problem of missing data among a set of predictor variables; and caution should be used in applying them in a situation where the criterion data are missing. For instance, methods to estimate how an individual would have responded to a questionnaire from background information available on a person may be justifiable, but to estimate criterion scores from the individual's background data is a questionable procedure.

### Method

Because the data set being considered in this paper was generated by dividing the total observation period into several intervals and computing a criterion tally within each interval, the resulting measures would be expected to correlate, thereby making the regression approach attractive. However, in order to avoid the computational difficulties that would be encountered when the amount of missing information varied from person to person, a modified method was developed in which the regression weights varied as a function of the amount of data available and the amount missing. Thus, a single equation can be used, regardless of the amount of missing data.

This modified regression procedure was compared to the deletion and prorating procedures that are often used when the criterion is treated as a single value for each person rather than as a set of repeated measures. Each method was used to adjust individual scores in a data set designed to simulate information obtained from individuals having varying degrees of exposure. The simulated data were generated by modifying actual data collected during a study of morbidity among U.S. Navy enlisted men. Starting with individuals who had complete data (and therefore equal amounts of exposure) to define the initial population, a series of sample data sets were generated using a monte carlo procedure to randomly delete portions of the illness data. Thus, the sample sets simulated data obtained from naval personnel who had different amounts of exposure as a result of transfers to and from the research setting, i.e., Navy ships. Using these data, the three methods for adjusting criterion data were compared in terms of the degree that estimates obtained with the modified criterion reproduced results obtained with the initial population data.

### Rationale for Adjusting Criterion Scores by Means of Linear Regression

Consider a study in which the sample is composed of two groups: Group 1 consists of individuals present for the entire study and Group 2 consists of individuals present for only the first half of the study. In this situation the data from Group 1 could be used to generate a regression equation for predicting criterion scores obtained during the second half of the study from the value obtained during the first half. This equation can be expressed as follows:

$$\hat{Y}_2 = W(Y_1 - \bar{Y}_1) + \bar{Y}_2 \quad [1]$$

where  $\hat{Y}_2$ ,  $Y_1$ , and  $W$  are the predicted score for the second half of the study, the actual score from the first half, and the regression weight, respectively.

The problem, of course, is that individuals are generally lost from a study at different points during the investigation. Therefore, a regression equation tailored to the circumstances of a particular individual probably would not fit the next person, and creating a separate equation for each person would be impractical at best. To develop a more practical method for predicting the criterion score corresponding to  $m_2$  observations from the score obtained during  $m_1$  observations, a method was developed to estimate  $\bar{Y}_{m_1}$ ,  $\bar{Y}_{m_2}$ , and  $W_{m_1 m_2}$  from  $\bar{Y}_1$ ,  $\bar{Y}_2$ , and  $W$ , respectively.

Reconsidering the data for Group 1, the criterion score measured during the first half of a study can be expressed as the separate criterion counts made during each observation period summed over the  $k$  observation periods in the first half of the study, i.e.,

$$Y_1 = \sum_{i=1}^k X_i \quad [2]$$

Similarly, the score for the second half of the study would be

$$Y_2 = \sum_{i=k+1}^{2k} X_i \quad [3]$$

If it is assumed that the criterion variance from one observation period to another is constant and that the correlation of the criterion counts from one time to the next is also constant, then it can be shown that

$$S_{Y_1}^2 = S_{Y_2}^2 = K S_x^2 [1 + r_{xx} (K-1)] \quad [4]$$

where  $S_x^2$  is the criterion variance during one observation period, and  $r_{xx}$  is the correlation between the criterion scores obtained on two different occasions.

In addition, when the above conditions are assumed, the covariance of  $Y_1$  and  $Y_2$  is

$$S_{Y_1 Y_2} = K^2 r_{xx} S_x^2 \quad [5]$$

Therefore,

$$r_{Y_1 Y_2} = (K^2 r_{xx} S_x^2) / \{K S_x^2 [1 + r_{xx} (K-1)]\} \quad [6]$$

$$= (K r_{xx}) / [1 + r_{xx} (K-1)] \quad [7]$$

The reader may recognize Equation 7 as the Spearman-Brown formula for the reliability of a lengthened test (cf. McNemar, eq. 12.20, 1969; Lord & Novick, eq. 5.10.1, 1968).

Solving Equation 7 for  $r_{xx}$  gives

$$r_{xx} = r_{y_1 y_2} / [K - r_{y_1 y_2} (K-1)] \quad [8]$$

Now, the above process will be reversed. However, this time, instead of using Equation 7, the formula for two tests of lengths  $m_1$  and  $m_2$  will be used (McNemar, eq. 12.17, 1969) to allow for the possibility of different lengths of exposure among individuals. This procedure yields the following equation for the correlation of the criterion counts made during  $m_1$  observations with those made during  $m_2$  observations:

$$r_{m_1 m_2} = r_{xx} / \{ [(1 + r_{xx} (m_1 - 1)) / m_1] [(1 + r_{xx} (m_2 - 1)) / m_2] \}^{1/2} \quad [9]$$

To find  $W_{m_1 m_2}$  where

$$W_{m_1 m_2} = r_{m_1 m_2} S_{m_2} / S_{m_1} \quad [10]$$

Equations 4 and 9 can be combined to yield

$$W_{m_1 m_2} = m_2 r_{xx} / [1 + r_{xx} (m_1 - 1)] \quad [11]$$

To obtain estimates of  $\bar{Y}_{m_1}$  and  $\bar{Y}_{m_2}$  it will be assumed that there is no overall group trend so that the criterion mean per sample per observation is

$$\bar{Y}_x = \bar{Y} / (m_1 + m_2) \quad [12]$$

where  $\bar{Y}$  is the mean criterion score for the entire study. Therefore, the mean criterion scores for  $m_1$  and  $m_2$  observations are

$$\bar{Y}_{m_1} = m_1 \bar{Y} / (m_1 + m_2) \quad [13]$$

and

$$\bar{Y}_{m_2} = m_2 \bar{Y} / (m_1 + m_2) \quad [14]$$

Finally, an estimate of the mean criterion score assessed during  $m_2$  observations ( $\hat{Y}_{m_2}$ ) from the score determined from  $m_1$  observations ( $Y_{m_1}$ ) can be derived by substituting  $W_{m_1 m_2}$  for  $W$ ,  $\hat{Y}_{m_1}$  for  $\hat{Y}_1$ , and  $\hat{Y}_{m_2}$  for  $\hat{Y}_2$  in Equation 1 to yield

$$\hat{Y}_{m_2} = \{ (m_2 r_{xx}) / [1 + r_{xx} (m_1 - 1)] \} \{ Y_{m_1} - [(m_1 \bar{Y}) / (m_1 + m_2)] \} + (m_2 \bar{Y}) / (m_1 + m_2) \quad [15]$$

Therefore, when some individuals have fewer criterion assessments than others, the deficiency can be estimated by first computing  $r_{y_1 y_2}$  and  $\bar{Y}$  for those people present during the entire study. Then, using Equation 8, an estimate of  $r_{xx}$  can be computed which, when used in combination with  $\bar{Y}$  in Equation 15, yields  $\hat{Y}_{m_2}$ —an estimate of the data not observed. Thus, the adjusted criterion score is

$$C = Y_{m_1} + \hat{Y}_{m_2} \quad [16]$$

i.e., the observed criterion score plus the estimate for the unobserved period.

## Data

Data obtained during a study of illness incidence aboard three U.S. Navy amphibious ships were used in the monte carlo analysis. These data were gathered during an overseas deployment which lasted approximately 6 months for each ship. Because of small differences in the length of deployment, only illness data obtained during the first 160 days were used. In addition to the illness data, three other variables were included. These consisted of job demands (1 = low physical demands, 2 = medium demands, and 3 = high physical demands), pay grade (pay level), and age. These variables were used to assess predictor-criterion correlations on the basis of previous studies (Pugh & Gunderson, 1979; Rahe, Gunderson, Pugh, Rubin, & Arthur, 1972), which found them to be significantly related to illness.

From a total of 1,241 individuals aboard the three ships (for at least one day of the deployment), those who had complete data on all of the above variables were selected for analysis. This resulted in a study population of 470 men. These data were then altered to simulate the effect of transferring crew members to or from a ship. The method of altering data for an individual involved two steps. First, a day from 2 to 159 was selected at random; and second, a random decision was made as to whether the day selected was to be the day the person transferred to or from the ship. If the day selected was to be the day "reported aboard," illness data prior to that day were removed; if the day selected was considered to be the "detached" day, the data after that day were deleted. Each time an individual's data were altered, a prorated criterion score ( $Y_p$ ) was created as follows:

$$Y_p = Y_{m_1} (m_1 + m_2) / m_1, \quad [17]$$

and a regression adjusted criterion score was derived using Equation 16.

To evaluate trends related to the amount of missing data, six different conditions were assessed. In the first condition 10% of the sample had some illness data removed, in the second 20% had data removed, in the third 30% had data removed, in the fourth 50%, in the fifth 70%, and in the sixth 90% of the individuals had some illness data removed. For each condition 30 samples were created by randomly selecting the designated percentage of individuals, creating the two adjusted criterion scores, then returning those individuals to the sample, and repeating the process. It is important to note that the above percentage for each condition refers to the proportion of individuals with missing data. With respect to the total amount of data, however, the percentages were about one-half of the above values because only a portion of each individual's data was altered. For example, under the first condition 10% of the persons in the sample had data altered, but only about 5% of all illness data were removed, since a portion of each person's data remained intact.

Analyses were designed to determine whether adjusting the illness data by prorating, regression, or simply deleting subjects with incomplete data generated statistics most like those found when complete data for all 470 individuals were used.

## Results

### Criterion Characteristics

The effect of adjusting the illness counts on estimates of the criterion mean and standard deviation are shown in Table 1. Inspection of the mean value of the 30 sample criterion mean scores for each method under each condition shows that all three methods yielded estimates near the population value. However, the last three columns in Table 1 show that the regression-derived values were most frequently the closest to the population value. Therefore, the regression methods appear somewhat more efficient than the other methods.

Table 1  
Means and Standard Deviations of Criterion Estimates for Each Condition

Data Condition	Coefficient	Means of Sample Values		Frequency of Being the Best (Closest) Estimate	
		Deletion	Regression	Deletion	Regression
0 <sup>a</sup>	Mean				
10	(3.234) <sup>b</sup>	3.240	3.236	7	18
20		3.228	3.224	8	18
30		3.228	3.236	5	12
50		3.241	3.238	9	14
70		3.298	3.239	2	18
90		3.143	3.195	3	9
0	Standard Deviation				
10	(2.720)	2.720	2.662	22	5
20		2.716	2.607	29	0
30		2.712	2.556	30	0
50		2.682	2.404	30	0
70		2.749	2.328	29	1
90		2.624	2.219	26	4

<sup>a</sup>percentage of population given missing data.

<sup>b</sup>population value shown in parentheses.

Comparing the three methods with regard to estimates of the criterion standard deviation reveals some divergent results. Regardless of the missing data condition, the deletion method produced the best estimates of the population value, becoming progressively smaller as the amount of missing data increased. Prorating produced estimates farthest from the population value. These estimates were consistently greater than the population value and increased at a relatively rapid rate as the amount of missing data increased.

Criterion scores generated by the regression and prorating methods were further evaluated by computing the amount of error introduced into the data. Error was assessed by subtracting the adjusted criterion score for an individual from the actual value, squaring the difference, summing the squared deviation over individuals, dividing the result by the number of individuals ( $N = 470$ ), and then taking the square root of the mean squared deviation. The results of this procedure are shown in Table 2. As would be expected, the mean error for both methods increased as the amount of missing data increased. However, the  $t$  values shown in Table 2 indicated that under every condition the amount of error introduced by prorating was significantly more than the amount due to the regression method.

Table 2  
Amount of Criterion Error Introduced  
by Regression and Prorating<sup>a</sup>

Data Condition <sup>b</sup>	Method				t
	Regression		Prorating		
	Mean	S.D.	Mean	S.D.	
10	.536	.007	1.114	.641	- 4.94
20	.735	.008	1.474	.211	-19.19
30	.893	.006	1.865	.739	- 7.20
50	1.173	.007	2.061	.176	-27.67
70	1.365	.006	2.592	.218	-30.86
90	1.566	.006	3.074	.407	-20.31

<sup>a</sup>Criterion error<sub>1</sub> is measured in terms of E for each sample,  $E = [\sum(y - \hat{y})^2 / N]^2$ ,  $y$  is the actual criterion score of the individual, and  $\hat{y}$  is the adjusted criterion score.

<sup>b</sup>Percent of population given missing data.  $N = 470$ .

### Predictor-Criterion Relationships

The correlations among the three predictor variables—job demands, pay grade, and age—and actual illness counts for the study population are shown in Table 3. Note that past illness did predict future illness in this population ( $r = .359$ ), that the correlations of job demands and pay with illness were significant, but that age did not correlate to a significant degree with illness.

To determine the effect that the various methods of dealing with missing criterion data had on correlational analyses, the correlations between the three predictor variables and the various criterion scores were computed. The 30 values of each coefficient were converted to  $z$  scores; the mean  $z$  score for each coefficient was determined and converted back to a correlation coefficient. These data are shown in Table 4. Several trends are notably similar for the three predictors. First, the regression and prorating methods consistently underestimated the population value, whereas the estimates obtained

Table 3  
Correlations among Actual Illness Counts  
and Three Predictor Measures

Variable	1	2	3	4	5
1. Job demands					
2. Pay grade	-.203*				
3. Age	-.229*	.737*			
4. Illness for days 1-80	.100*	-.094*	-.017		
5. Illness for days 81-160	.092*	-.133*	-.085	.359*	
6. Illness for days 1-160	.116*	-.139*	-.065	.797*	.850*

\* $p < .05$ ;  $N = 470$

through deletion were above and below the population value, depending upon the level of missing data. Second, the mean predictor-criterion correlation estimates from both regression and prorating grew smaller as the amount of missing data increased, with prorating declining at a faster rate. Third, mean estimates generated by the regression method were consistently closer to the population value than the mean estimates produced by the prorating method. Finally, the regression method appeared most efficient, producing estimates closer to the population value than either deletion or prorating in 17 out of 18 cases. The one occasion where regression was not closest was when age was being correlated with the criterion scores under the 20% missing data condition and then prorating was most often closest.

In order to summarize the above data and to illustrate the salient properties of each method, multiple correlations of the three predictors with each criterion were computed. Means and standard deviations for these multiple correlations were computed for each method under each condition, and the results are depicted in Figure 1. The dotted lines indicate the critical value for a multiple correlation at the .01 probability level. It can be seen that as the amount of missing data increased (1) the values generated by the deletion method tended to vary about the population mean until extreme levels of missing data were reached, (2) the values generated by the regression and prorating methods showed a steady decline with means for prorating having the steeper gradient, and (3) the regression values which had the smallest amount of variation remained above the critical level ( $\alpha = .01$ ) more often than the values generated by either the deletion or prorating methods.

### Discussion and Conclusions

Viewing the criterion comparability problem as a missing data problem appears to be useful because the results obtained in the monte carlo analyses indicated that the procedure designed to estimate missing information, the regression procedure, compared favorably with the deletion and prorating procedures. In particular, with respect to predictor-criterion correlations, the regression method appeared to be more efficient than either of the other methods because the values obtained with the regression method most often were the best estimates of the population coefficient. Therefore, in a situation where data have been collected for correlational analysis and a portion of the criterion information was not observed for some individuals, the regression method may be a more attractive manipulation than that provided by either prorating or deletion. Although the correlations

Table 4  
 Estimates of Predictor-Criterion Correlations for Each Condition

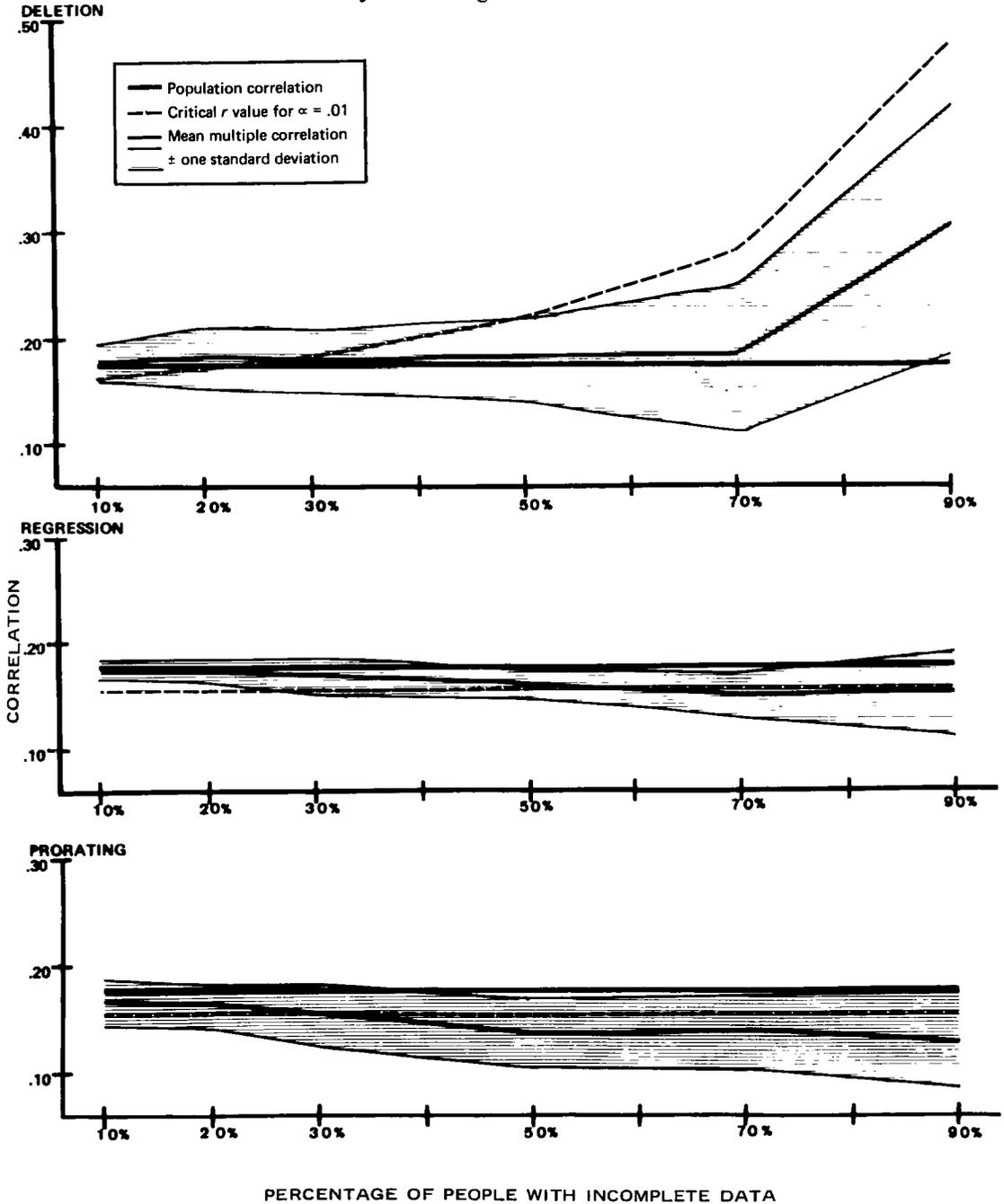
Data Condition	Means of Sample Values		Frequency of Being the Best (Closest) Estimate	
	Deletion	Regression Prorating	Deletion	Regression Prorating
0 <sup>b</sup>				
10	.119	.115	8	17
20	.122	.114	6	14
30	.110	.106	8	17
50	.122	.104	6	18
70	.102	.095	7	21
90	.188	.099	3	22
0				
10	-.138	-.135	4	18
20	-.136	-.133	6	16
30	-.139	-.132	9	17
50	-.136	-.121	10	15
70	-.103	-.110	4	17
90	-.147	-.108	3	16
0				
10	-.066	-.064	8	17
20	-.063	-.061	5	12
30	-.062	-.061	9	14
50	-.072	-.055	5	21
70	-.035	-.047	7	13
90	-.068	-.047	4	11
0				
10	-.065	-.059	8	17
20	-.061	-.051	5	13
30	-.062	-.050	9	14
50	-.072	-.038	5	21
70	-.035	-.040	7	13
90	-.068	-.034	4	11

<sup>a</sup>x<sub>1</sub> = Job Demands; x<sub>2</sub> = Pay Grade; x<sub>3</sub> = Age

<sup>b</sup>Percentage of population given missing data.

<sup>c</sup>Population value shown in parentheses.

**Figure 1**  
 The Mean and Standard Deviation of Sample Multiple Correlations Obtained After Missing Criterion Data Were Adjusted Using Each of Three Methods



obtained tended to underestimate slightly the value that would be obtained from complete data, they would be expected to be closer than the coefficients generated by either of the other methods. Further, by retaining individuals in the analysis, enough power can be gained to more than compensate for the criterion bias. However, the regression estimates are biased; and if circumstances are such that it is more important for the estimator to be unbiased than efficient, then the deletion method would be preferred.

The above conclusions should be qualified, however. Because individuals were selected at random to receive missing data, caution should be used in extending the results to situations where the lack of data occurs in a systematic fashion. Moreover, many assumptions about the homogeneity of variances, covariances, and means were made when deriving the regression formulas. To the degree that a data set fails to meet these assumptions, it can be expected that the performance of the regression method will degenerate. Nevertheless, reservations about the performance of the regression method in a "real" study may be relieved somewhat by noting that the original data set (i.e., population data) was gathered for a study of illness among Navy men. In this way many of the deviations that actual data take from idealized states were represented in the monte carlo analyses. Finally, the present interpretations depend upon the assumption that estimating what happened when a person was present, but not observed, is equivalent to predicting what would have happened if the person had been present when, in fact, he/she was not.

With due regard for the above points, it is felt that the regression method can be a useful technique for adjusting criterion data when some individuals have less exposure time than others. It should be pointed out that the regression technique requires a sufficient number of individuals with complete data in order to estimate the degree to which past behavior is associated with future behavior. In addition, individuals with incomplete data should have enough actual information to use as a basis for estimating the missing information. If it is found that many individuals with missing data have a minimal amount of criterion data (i.e., not enough data to project from), little would be gained by retaining them in the analysis. In such circumstances, the deletion method may be preferred. Of course, when there are excessive amounts of missing information—large numbers of individuals with large amounts of missing data points—no method of analysis would be advisable.

Although illness data have been used to illustrate and evaluate the various procedures, researchers studying other criteria may also find the regression method useful. For example, studies of worker performance in an industrial setting might employ the regression method for analysis of data retained from individuals who transfer to a new company. This may be particularly desirable if these individuals had participated in an extensive pretesting and/or training program prior to the collection of performance data. Before using the regression method, however, an investigator may wish to inspect the split-half criterion correlation. As this correlation becomes unity, Equation 16 reduces to Equation 17, and prorating becomes equivalent to regression. Therefore, if the split-half criterion correlation is found to be very high, the available criterion data can simply be prorated. On the other hand, if the split-half criterion correlation approaches zero, the regression slope becomes zero and the missing data are replaced by a prorated mean score. However, if it is found that the split-half correlation is in fact near zero, then the criterion measures would have no reliability—a most challenging object for study. Finally, in actual practice the researcher should consider trying all the methods and comparing the results obtained from each one.

In addition, when evaluating the usefulness of the regression procedure for a particular problem, decisions regarding (1) the adequacy of the split-half criterion correlation, (2) the randomness of missing data, (3) the amount of missing data per person and per sample, and (4) the homogeneity of means, standard deviations, and correlations must be made in the context of the costs of various types

of errors. One must consider, for example, if the amount of bias introduced by adjusting the data causes an important and valid hypothesis to be rejected. Perhaps, instead of adjusting the criterion, more data should be collected. Also to be considered is the cost of the wasted data if individuals are dropped from an analysis. The procedures, analyses, and results of the present paper were designed to aid the researcher in answering such questions. Because the circumstances of each study are different, the answers will tend to vary.

### References

- Buck, S. F. A method of estimation of missing values in multivariate data suitable for use with an electronic computer. *Journal of the Royal Statistical Society, Series B*, 1960, 22, 302-306.
- Edgett, G. L. Multiple regression with missing observations among the independent variables. *Journal of the American Statistical Association*, 1956, 51, 122-131.
- Frane, G. W. Some simple procedures for handling missing data in multivariate analysis. *Psychometrika*, 1976, 41, 409-415.
- Galdi, J., & Bonato, R. R. Biasing effects of missing data on a family illness criterion. *Psychological Reports*, 1977, 41, 279-288.
- Glasser, M. Linear regression analysis with missing observations among the independent variables. *Journal of the American Statistical Association*, 1964, 59, 834-844.
- Gleason, T. C., & Staelin, R. A proposal for handling missing data. *Psychometrika*, 1975, 41, 229-252.
- Lord, F. M., & Novick, M. R. *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley, 1968.
- McNemar, Q. *Psychological statistics* (4th ed.). New York: John Wiley & Sons, 1969.
- Pugh, W. M., & Gunderson, E. K. E. Environmental factors in the onset of illness aboard Navy ships (Report No. 79-4). San Diego, CA: Naval Health Research Center, 1979.
- Rahe, R. H., Gunderson, E. K. E., Pugh, W. M., Rubin, R. T., & Arthur, R. J. Illness prediction studies: Use of psychological and occupational characteristics as predictors. *Archives of Environmental Health*, 1972, 25, 192-197.
- Rubin, D. B. Formalizing subjective notions about the effect of nonrespondents in sample surveys. *Journal of the American Statistical Association*, 1977, 72, 538-543.
- Timm, N. H. The estimation of variance-covariance and correlations matrices from incomplete data. *Psychometrika*, 1970, 35, 417-437.

### Acknowledgments

*This paper was Report Number 79-51, supported by Naval Medical Research and Development Command, Department of the Navy, under Research Work Unit MF58.524.023-0011. The views presented in this paper are those of the author. No endorsement by the Department of the Navy has been given or should be inferred. The author acknowledges the assistance of Ms. Mary Paul, who with enthusiasm and dedication, provided the computer support required for the data analyses reported on in this paper and to Ms. Pat Polak for her preparation of the manuscript.*

### Author's Address

Send requests for reprints or further information to William M. Pugh, Naval Health Research Center, San Diego, CA 92138.