Public Health Briefs

Reliability of Two Measures of Life Stress Among Outpatients at a Veterans Hospital

GLORIA A. BAILEY, MSPH, THOMAS D. KOEPSELL, MD, MPH, AND DONALD W. BELCHER, MD

Abstract: To measure test-retest reliability, 38 Veterans Administration clinic patients completed a modification of the Sarason Life Experience Survey (LES) twice within several weeks. Summary scores were found to be more reliable than scores on individual questions. Patients attending the Mental Hygiene Clinic responded less reliably than other patients. The "stress ladder"—a simple summary score measure—was found to be the most reliable measure of all. (*Am J Public Health* 1984; 74:723–724.)

Introduction

A clinical tool to assess the occurrence of certain "life events" as a reflection of stress was described by Meyer over 30 years ago.[1] The well known Holmes-Rahe Social Readjustment Rating Scale (SRRS) is the best known currently used instrument.[2] More recent adaptations of this scale have addressed how recently a life event occurred and whether the respondent regarded its impact as positive or negative.

Horowitz evaluated the reliability of an adaptation of this scale using 27 psychiatric outpatients. In a six-week test-retest study, he found the correlation between the number of items checked both times was .82. However, for specific times, there was only 60 per cent concordance between reported occurrence on the initial administration and that reported on the second administration.[3]

Sarason determined the reliability of his Life Experience Survey (LES) with students by giving the same respondent another LES after five to six weeks.[4] While the positive events score was found to be unreliable ($r = .19$ in one test and $r = .53$ in another), the negative and total scores were moderately reliable at .56 and .88 for the negative scale correlation coefficient and .63 and .64 for the total score correlation.

The present study was undertaken to test the reliability of an adaptation of a questionnaire on patients in a Veterans Administration clinic population before using it in a subsequent study. While the scale had been tested previously on college students, veterans do not have as much test-taking experience as students and might not respond the same way. It was also desirable to determine if there was another measure of stress that would be easier for veterans to use and not as time consuming.

Methods

The Life Events Questionnaire used in this study was adapted from that developed by Sarason.[4] Several events were omitted and several were added for a total of 46 items, to make the questionnaire more applicable to the mostly blue collar veteran population.*

In order to have a simpler measure of stress, we appended a 20-rung "stress ladder" at the end of the Life Events Questionnaire to allow respondents to self-rate their stress level from 1 to 20. The stress ladder was first used by Kirsh.** Patients are asked to state which numbered rung on a ladder best indicates how much stress they have experienced during the past year.

Subjects for this study were drawn from patients attending the Health Promotion Clinic (HPC) at the Seattle Veterans Administration Medical Center. At the time of each patient's visit with the HPC nurse practitioner, he/she was given a self-administered stress questionnaire to complete. Patients with vision, reading, or other problems were helped by a research assistant.

About 10 days after the initial questionnaire was completed, the first 49 patients were mailed a second identical questionnaire. Thirty-eight patients (77.6 per cent) returned the second questionnaire.

Nine of the 38 patients who responded had been patients in the Mental Hygiene Clinic (MHC) at some time during the previous two years. They were analyzed separately from the other 29 because of the previously reported lower reliability of such patients.[5]

Kappa coefficients[6] were calculated for each specific event. In addition, intraclass correlation coefficients were calculated for various summary statistics indicating test-retest reliability.[7] Pearson correlation coefficients were also calculated to test how well each summary statistic correlated with the stress ladder score.

Summary scores were calculated in two ways: a non-weighted score was calculated in which each event counted

Address reprint requests to Gloria A. Bailey, MSPH (doctoral candidate), Department of Epidemiology, University of Washington, Seattle, WA 98195. Dr. Koepsell is also with the Department of Epidemiology; Dr. Belcher is with the Department of Medicine, both at the University of Washington. All authors are also affiliated with the Northwest Health Services Research and Development Program, Seattle Veterans Administration Medical Center. This paper, submitted to the Journal July 26, 1983, was revised and accepted for publication January 9, 1984.

*Details available upon request to the authors.

**Kirsch EB: Psychological stress and the pathogenesis of genital herpes infections. Unpublished thesis (MPH), University of Washington, 1978.

equally; weighted scores were calculated by multiplying each event by the respondent's rating of the importance of that event (no effect = 0, a little = 1, some or a moderate amount = 2, and a great deal = 3). Positive events (rated as good by the patient) and negative events (rated as bad) were scored and tested separately as well as combined.

Results

Kappa coefficients were calculated for each specific event.* Summary scores are listed in Table 1. Positive scores were somewhat more reliable than negative scores. However, the composite positive score minus negative score was the most reliable of all summary scores. As expected patients with disorders treated in a mental hygiene clinic reported less reliably than others.

According to Fleiss,[6] a Kappa score of greater than .75 or so may be taken to represent excellent agreement beyond chance, values below about .40 may be taken to represent poor agreement beyond chance, and values between .40 and .75 may be taken to represent fair to good agreement beyond chance. Only 20 of the 41 events in the LES had a Kappa score over .40.*

TABLE 1—Reliability of Life Events Scores

Summary Score	Intraclass Correlation Coefficient between Scores on First and Second Questionnaires		Pearson r between Score and Ladder on First Questionnaire
	non MHC** (n = 29)	MHC (n = 9)	(n = 38)
Total number of negative events	.422*	.418*	.386*
Total number of positive events	.610*	.132	.083
Total number of events (positive and negative)	.395*	.062	.301*
Weighted negative events score	.431*	.418	.467*
Weighted positive events score	.602*	.313	.029
Positive score − negative score	.728*	.670	−.338*
Positive score + negative score	.318	.074	.367*
Stress Ladder	.863*	.954*	—

*$p < .05$
**MHC denotes patients who had attended the Mental Hygiene Clinic at any time during the past two years.

*Details available upon request to the author.

Discussion

One factor in the data collection protocol for the study which may have affected the results is that the test-retest questionnaires were not taken under exactly the same circumstances both times. The patient and his spouse may have completed the retest questionnaire together in some cases.

The reliability of the summary scores is greater than consideration of individual Kappa scores might indicate. People with high scores initially still score high on repeated administration, even though different events might be indicated.

The stress ladder is of considerable interest because of its high reliability among MHC as well as other patients. There is a reasonable correlation between the negative life events score and the stress ladder score, while there is no correlation between the positive score and the ladder score. The validity of the stress ladder as a predictor of future events is yet to be shown, but its higher reliability and simplicity of application and interpretation make it a promising candidate as a measure of stress.

It appears to be highly desirable to test the reliability of Life Events Questionnaires among the population who will be using them rather than simply relying on published values from other populations.

REFERENCES

1. The Collected Papers of Adolf Meyer. *In:* Winters EE (ed): Medical Teaching, Vol III. Baltimore: Johns Hopkins Press, 1951; 52–56.
2. Holmnes TH, Rahe RH: The social readjustment rating scales. J Psychom Res 1967; 11:213–218.
3. Horowitz M, Schaffer C, Hiroto D, Wilner N, Levin B: Life event questionnaires for measuring presumptive stress. Psychosom Med 1977; 39:413–439.
4. Sarason IG, Johnson JH, Siegel JM: Assessing the impact of life changes: Development of the life experiences survey. J Consult Clin Psychol 1978; 46:932–946.
5. Lawton AM, Seabury C, Branch LC, Azar GJ, Bond JO: Follow-up studies of St. Louis encephalitis in Florida. South Med J 1966; 59:1409.
6. Fleiss JL: Statistical Methods for Rates and Proportions. New York: John Wiley and Sons, 1981.
7. Winer BJ: Statistical Principles in Experimental Design, 2nd Ed. New York: McGraw-Hill, 1971.

ACKNOWLEDGMENTS

The authors wish to thank the Research Assistants, Mary Kay Morgan and Marcia Hunt, as well as the HPC Nurse Practitioners, Nancy Risser, Barbara Sullivan, and Lori Taylor, for their assistance in the administration of the questionnaires. We also wish to acknowledge Dr. Edward Kirsch who first devised and used the stress ladder. Partial support for this study was provided by the Health Services Research and Development Program of the Veterans Administration.