Chest Muscle Activity and Panic Anxiety: A Preliminary Investigation

PATRICK LYNCH, PHD, DONALD A. BAKAL, PHD,
WILLIAM WHITELAW, MD, AND TAK FUNG, PHD

This report represents a pilot investigation of the role of chest muscle electromyographic (EMG) activity in developing panic episodes. Chest EMG activity was obtained as part of a larger study examining ventilatory differences between panic sufferers and normal controls. Frontalis EMG, heart rate, and minute ventilation (breathing rate and tidal volume) were also obtained during the study. The ventilatory procedure involved exposing the subjects to three periods of carbon dioxide gas inhalations (1%, 3%, 5%; balance oxygen). Subjective measures of frightening cognitions and body sensations were obtained across the inhalation phases as well. The panic disorder subjects were divided, on the basis of subjective anxiety ratings obtained throughout the study, into high anxious (HA) and low anxious (LA) panic disorder groups. The HA panic disorder patients exhibited significantly higher chest EMG activity than the LA panic disorder patients and controls across all phases of the experiment. In addition, the chest EMG predicted, better than the other physiologic measures, the number of frightening cognitions and sensations reported by the subjects during the baseline and 5% CO_2 inhalation phases. Overall, the results were supportive of the further study of chest wall EMG activity in the pathogenesis of panic attacks.

INTRODUCTION

The multifaceted origins of panic disorder are recognized to include both cognitive and physiological determinants, with hypotheses surrounding central biological panic sites continuing to dominate the literature (1). There is growing appreciation, however, that factors outside the central nervous system per se contribute to both susceptibility and onset of panic attacks. Psychological or cognitive variables are recognized as part of the puzzle (2, 3) but purely psychological formulations have difficulty explaining the powerful physical symptoms of panic as well as accounting for attacks in the absence of provocation.

This study examined chest wall muscle activity as part of an investigation dealing with ventilatory correlates of panic susceptibility. The larger study examined two chronic hyperventilation hypotheses based on chemoreceptor sensitivity and respiratory after-discharge (4). The ventilatory hypotheses were tested using different levels of CO_2 inhalation and guided by the expectation that panic disorder patients would evidence greater increases than controls in minute ventilation in response to 1%, 3%, and 5% CO_2 concentrations. Chest electromyographic (EMG) activity was studied as a possible physiologic component of panic reactions to the experimental situation and manipulations.

From the Division of Psychology, Foothills Hospital (P.L.); and the Department of Psychology (D.A.B.), Respiratory Medicine (W.W.), and Academic Computing Services (T.F.), University of Calgary, Calgary, Alberta, Canada.

Address reprint requests to: Donald A. Bakal, 2500 University Drive N.W., Calgary, Alberta, T2N 1N4.

Received September 6, 1990; revision received October 2, 1990

0033-3174/91/5301-0080$03.00/0

The decision to monitor chest wall muscle activity during the provocation phases of the experiment was made on the basis of symptoms commonly reported by panic sufferers. Dyspnea is the most common respiratory symptom reported to precede and accompany panic episodes (2). However, the majority of panic sufferers also describe feelings of pressure, tightness, heaviness, and other sensations from the chest. Recent clinical observations of panic sufferers attest to the pervasive presence of chest pain as well as chest pressure and tightness prior to and during panic episodes (2, 4). The physiological origins of chest pain in panic disorder patients are unknown; however, cardiovascular disease does not appear to be implicated. Research has shown a high incidence of the diagnosis of panic disorder and normal coronary arteriograms in chest-pain patients presenting to cardiologists (5).

Chest wall muscle activity as reflected in electromyographic activity (EMG) from the region of the sternum was recorded across all phases of the experiment for 9 panic disorder patients and eight normal controls. Forehead EMG, heart rate, and minute ventilation were also monitored. Subjective measures of anxiety, frightening cognitions, and somatic symptoms were also obtained across all phases in order to examine the interplay between psychological and physiological components of panic.

SETTING AND SUBJECTS

The study was conducted at the Pulmonary Laboratory/Clinic of the Foothills Hospital, Calgary, Alberta. Female psychiatric outpatients making their initial presentation for treatment were interviewed by a clinical psychologist who, as part of a general diagnostic interview, applied DSM-III (6) criteria. Nine subjects with a primary diagnosis of panic disorder (with or without agoraphobia), who reported at least four of the 12 DSM-III symptoms to be moderately severe during episodes of anxiety, and a frequency of panic at least once per week, were selected for the study. Eight female normal controls were selected from hospital and university staff. Questionnaire assessment of panic symptoms was used as a validity check of the diagnosis. Each subject completed the Body Sensations Questionnaire (BSQ) (7), the Agoraphobic Cognitions Questionnaire (ACQ) (7), and the Fear Questionnaire (8). The ACQ consists of 14 items dealing with thoughts of physical catastrophe due to anxiety symptoms and with thoughts reflecting social or behavioral disaster from loss of control. The BSQ is a 17-item scale concerning the degree to which patients fear somatic symptoms commonly associated with panic. Total Phobia Score and Agoraphobia Score were selected from the Fear Questionnaire. The same clinical interview and questionnaires were given to female control subjects selected from a group of hospital and university staff volunteers. Only females were used as an attempt to reduce variability on the baseline pulmonary function measures.

All patients and controls received a brief medical examination to exclude subjects who might have organic respiratory or cardiac disease or hyperthyroidism. All subjects provided informed consent. No subjects withdrew during the course of the experiment.

METHOD

Physiologic Measures

The primary physiologic measures relevant to this report are chest electromyogram (EMG), frontalis EMG, minute ventilation, and electrocardiogram (EKG). Chest and forehead EMG measures were each obtained from sets of three Beckman silver/silver chloride electrodes. The skin was prepared with prepping paste and isopropyl alcohol and the electrodes were attached 1.0 cm apart by means of an electrode adhesive strip. The three chest electrodes were placed vertically on the sternum, with the lowest 5 cm above the xiphoid process. The EMG measures were obtained from Cyborg EMG J33 preamplifiers and a Cyborg dual EMG rms contour processing system (100–1000 Hz passband, time constant = 0.3 seconds). In order to minimize artifact

due to changes in the breathing cycle and the EKG, the lowest EMG recorded during the last 10 seconds of each minute of observation was used as the recorded value. This was an end-expiratory value. EKG was recorded using a five lead (left leg, right leg, left arm, right arm, chest) configuration, with the signals routed through a Hewlett-Packard amplifier. Expired air was continuously sampled from the mouthpiece through which the subject breathed. End tidal CO_2 was monitored using a Beckman LB-II Infra-Red Gas Analyzer, which was calibrated before data collection using room air and a calibration-grade medical gas mixture containing 5.02% carbon dioxide, 94.98% oxygen. Minute ventilation was determined on the basis of expired flow, which was measured by a Fleisch #2 pneumotachograph and a Validyne differential pressure transducer attached to the output side of an Otis-McKerrow valve, with the signal sent to a Hewlett-Packard Carrier amplifier. The gas mixtures were administered from a rubber reservoir bag connected to the input side of the Otis-McKerrow valve. The reservoir was kept partially full of gas through a continuous flow from the calibrated gas cylinder. With each change in gas concentration, the reservoir and connecting tubes were thoroughly flushed and filled with the next gas concentration.

Procedure

Prior to the laboratory session, subjects were informed that this was a study of respiratory parameters, and that they would be exposed to a series of CO_2/O_2 gases. They were informed that any changes in breathing during the inhalations would be a normal consequence of the gas inhalations. During the laboratory session, subjects were required to lie supine and breathe different mixtures of CO_2-O_2. In order to reduce anxiety, patients were informed in a session one week before the lab session and during the lab session that this was not a study of panic attacks, but of their general physiology, and were reassured that measures would be taken to reduce the possibility that they would panic in the laboratory. Patients and controls were informed in advance that inhalation of various mixtures of CO_2 might cause some physical symptoms similar to those associated with physical exercise (e.g., higher respiration rate, fatigue).

Each session began with the measurement of the subject's Forced Vital Capacity (FVC) and Forced Expiratory Volume in one second (FEV1) while standing and breathing into a spirometer. Next, the subject was required to lie on a bed while the EMG and EKG electrodes were attached. The subject was then familiarized with the use of the mouthpiece, following which a noseclip was affixed so that all breathing would be entirely through the mouthpiece. Once everything was in place, the five-minute baseline period began during which time subjects were instructed merely to lie quietly and get used to the equipment while breathing room air. At the end of this period, the noseclip and mouthpiece were removed and the rating scales were completed. The baseline was followed by three successive five-minute CO_2 gas inhalation periods with 1%, 3%, and 5% CO_2 in oxygen, respectively. The subjective rating scales were completed at the end of each inhalation period. Calibration grade CO_2 gas mixtures were used throughout the study. The 5% CO_2 inhalation was followed by a five-minute recovery period.

Immediately after each of the five phases, the subject was prompted to rate her level of anxiety on a 10-point Likert scale ranging from '0' (very relaxed and sleepy) through '5' (moderately severe anxiety) to '10' (full-blown panic attack). Responses to this scale were used to group subjects into those who showed evidence of panic anxiety across the manipulations from those who did not. Subjects were also required to report on the intensity of a number of somatic and cognitive symptoms as measured by the BSQ and the ACQ. The instructional set given to subjects to complete these instruments was in reference to their experiences during the immediately preceding phase.

RESULTS

Pre-Baseline Comparisons

The clinical and control subjects did not differ on age, weight or height (Table 1). The groups were also similar in terms of forced vital capacity and forced expiratory volume. The groups were significantly different in terms of their scores on the Agoraphobic Cognitions Questionnaire (ACQ; $t = 4.03$, $df = 15$, $p < 0.002$) and the Body Sensations Questionnaire (BSQ; $t = 7.19$, $df = 15$, $p < 0.001$). The panic disorder subjects also reported significantly more phobic avoidance behavior than the controls as measured by the

TABLE 1. Comparison of Panic Disorder Patients and Controls on Pre-Experimental Physical and Psychological Characteristics*

	Panic Disorder		Controls	
	M	SD	M	SD
Age (years)	36.4	9.7	31.1	8.3
Weight (kg)	64.5	16.2	61.6	12.4
Height (cm)	163.6	6.3	161.8	7.3
FEV1 (liters)	3.6	0.5	3.5	0.6
FVC (liters)	4.2	0.6	4.1	0.6
Cognitions	35.9	11.4	17.8	3.2
Sensations	50.7	10.1	23.0	6.7
Total phobia	53.4	14.9	10.9	5.5
Agoraphobia	20.2	1.9	10.3	1.6
No. of panics	4.9	3.1	0.0	0.0

* FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity; Cognitions, total score on Agoraphobic Cognitions Questionnaire (ACQ); Sensations, total score on Body Sensations Questionnaire (BSQ); Total Phobia, subscale scores from the Fear and Agoraphobia Questionnaire; No. of panics, number of panic attacks reported during last 7 days.

Total Phobia score ($t = 7.5$, $df = 15$, $p < 0.001$) and the Agoraphobia scale ($t = 4.9$, $df = 15$, $p < 0.001$) of the Fear Questionnaire.

Each subject was queried about medication use during the 24 hours prior to the laboratory session. One panic disorder subject reported minor tranquilizer use and two subjects reported tricyclic antidepressant use. None of the panic subjects reported use of monoamine oxidase inhibitors or beta-blockers. None of the controls reported use of any medications.

Sub-Grouping of Panic Disorder Subjects

Although efforts were made to minimize the occurrence of anticipatory anxiety and panic, the subjects proved to be highly variable in their subjective reactions to the laboratory situation, both at baseline and across the various CO_2 inhalation phases. Prior to the analyses, the subjects were grouped on the basis of how "anxious" or "panicky" they became during the various phases of the experiment. The grouping was made on the basis of each subject's average response across the five administrations of the 10-point subjective anxiety scale. Previous research (9) demonstrated that the average panic attack occurs at a moderate intensity (5.6 on a 0–10 scale). Subjects with an average scale score of 5 or greater ($n = 6$) were placed in the High Anxious (HA) group while subjects with an average score less than 5 were placed in the Low Anxious (LA) group ($n = 3$). None of the control subjects ($n = 8$) obtained an average anxiety score above 5. All three panic disorder subjects reporting medication use were in the HA group. An additional analysis using the Marks and Mathews (8) Fear Questionnaire indicated that the three groups also differed in phobic severity outside the experimental situation (Table 2). HA subjects had significantly greater Total Phobia scores, $F(1,14) = 6.96$, $p < 0.02$, and significantly greater Agoraphobia scores, $F(1,14) = 7.63$, $p < 0.02$, than the LA group. The LA group also had significantly higher scores on the Total Phobia subscale, $F(1,14) = 21.35$, $p < 0.001$, and the Agoraphobia subscale,

TABLE 2. Group Means for the total Phobia and Agoraphobia Subscales of the Fear Questionnaire

	Total Phobia		Agoraphobia	
	M	SD	M	SD
HA panic disorder	59.5	13.2	24.3	9.0
LA panic disorder	41.3	11.0	12.0	8.2
Controls	10.9	5.5	1.9	1.6

$F(1,14) = 5.61$, $p < 0.05$, than the Control group.

Physiologic Responses to Carbon Dioxide

The physiologic data were grouped and analyzed across the baseline, CO_2 inhalation, and recovery phases. The data were reduced to one-minute averages for the time periods defining each phase. The group means for chest EMG, forehead EMG, heart rate and minute ventilation are presented in Figure 1. The primary variable of interest in this study was chest EMG activity. However, given the potential interdependence of the physiological variables, a three-way repeated measures MANOVA (Groups, Phases, Time Within Phase) was carried out using a total of six physiological measures (breathing rate and tidal volume were included as variables separate from minute ventilation). The multivariate effects for Phase, $F(24,179) = 4.91$, $p < 0.01$, and Time Within Phase, $F(24,179) = 2.83$, $p < 0.01$, were significant. The univariate F's for the three ventilatory measures were significant for both the Phase and Time effects ($p < 0.01$) indicating that the CO_2 manipulations were effective in altering breathing. Heart rate also showed a significant univariate F for Phases, $F = 2,31$, $p < 0.001$, but not for Time. Chest and forehead EMG however were not associated with significant univariate F's for Phases or Time.

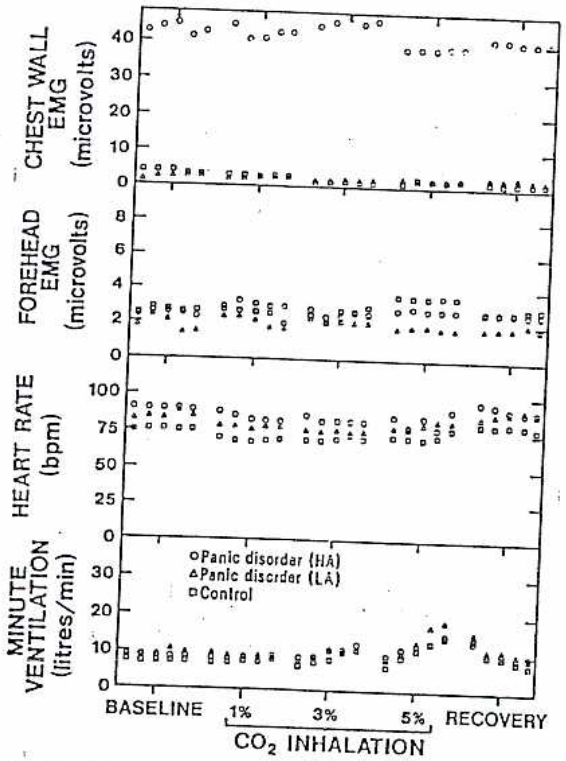

Fig. 1. Physiologic average scores of high anxious (HA) panic disorder subjects, low anxious (LA) panic disorder subjects, and controls.

The multivariate F for Groups was nonsignificant. The a priori hypothesis regarding chest EMG was supported with a significant univariate result, $F(2,14) = 6.58$, $p < 0.01$. The Group univariate F's for the remaining physiological variables were all nonsignificant. Post-hoc analyses revealed that the HA panic group had significantly higher chest EMG than both the LA panic, $F(1,14) = 6.83$, $p < 0.02$, and control groups, $F(1,14) = 11.58$, $p < 0.01$. The LA group did not differ significantly from the control group on this measure. In summary, heightened chest muscle activity was observed only in the panic disorder subjects who were also showing evidence of anxiety and/or panic. The group differences were independent of the CO_2 manipulations.

As a further examination of possible ventilatory differences, a separate one-way ANOVA was conducted on the final minute of each of the three CO_2 inhalation phases. This minute represents the period of maximal provocation within

each gas inhalation phase and the best approximation of steady-state responding achieved in the study. None of the comparisons was significant.

Psychologic Responses to Carbon Dioxide

The modified BSQ and ACQ questionnaires were administered at the end of the baseline and immediately after each of the three carbon dioxide inhalation phases. A two way (Groups, Phases) MANOVA performed on the total scores of these questionnaires revealed significant main effects for Groups ($F(4,26) = 2.75$, $p < 0.05$) and for Phases ($F(6,82) = 2.42$, $p < 0.05$). Insufficient degrees of freedom ruled out calculation of the interaction effects. The Group univariate F for BSQ scores was significant, $F(2,14) = 5.45$, $p < 0.02$, while the Group univariate F for ACQ scores was nonsignificant. Post-hoc analyses revealed significant BSQ score differences between the HA Panic Group and Controls, $F(1,13) = 3.39$, $p < 0.01$. The remaining comparisons were nonsignificant.

The Phase univariate F for BSQ scores was significant, $F(2,32) = 3.35$, $p < 0.05$) while the Phase univariate F for ACQ scores was nonsignificant. As shown in Figure 2, the three groups showed a decrease or no change in their BSQ scores from baseline through to 3% CO_2 inhalation. Following 5% CO_2 inhalation, the groups showed an increase in their BSQ scores.

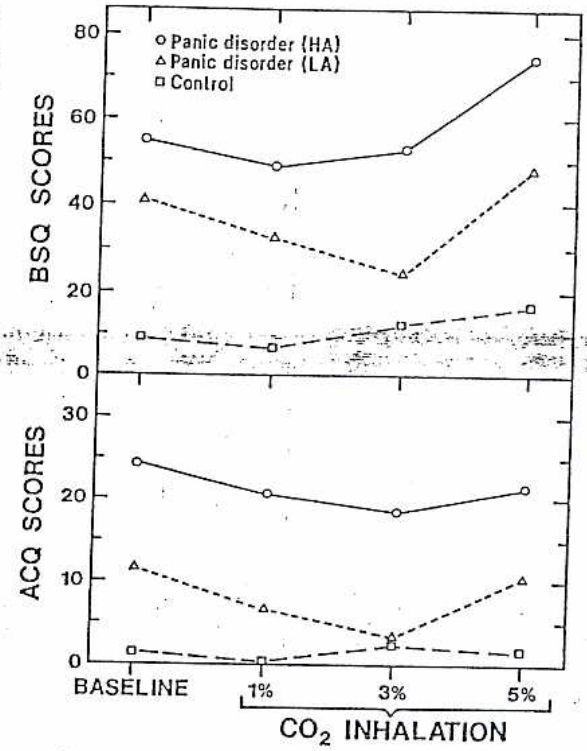

Fig. 2. Body Sensation Questionnaire (BSQ) and Agoraphobic Cognition Questionnaire (ACQ) scores obtained at the end of baseline and following each of the carbon dioxide inhalation phases.

Physiologic Predictors of Anxiety, Cognitions, and Bodily Sensations

The relationships between the physiologic and subjective measures during baseline and 5% CO_2 were examined with regression analyses. For each phase, a separate analysis was performed using BSQ scores and ACQ scores as the dependent variables, respectively. The independent variables were chest EMG, forehead EMG, minute ventilation, and heart rate. The physiologic data used in these analyses were obtained from the last minute of baseline and of 5% CO_2 inhalation. The results of these analyses are presented in Tables 3 and 4. The combination of chest EMG, forehead EMG, minute ventilation, and heart rate yielded significant R2 for

TABLE 3. Physiologic Predictors of BSQ and ACQ Scores During Baseline

Independent Variables	Dependent Variables	
	BSQ	ACQ
	Regression Coefficient	
Chest EMG	0.55*	0.76†
Forehead EMG	1.59	0.28
Minute vent	4.96*	2.31*
Heart rate	0.92*	0.26

* $p < 0.05$.
† $p < 0.001$.

TABLE 4. Physiologic Predictors of BSQ and ACQ Scores Following 5% CO2 Inhalation

Independent Variables	Dependent Variables	
	BSQ	ACQ
	Regression Coefficient	
Chest EMG	1.26‡	0.88†
Forehead EMG	0.54	−0.36
Minute vent	0.24	1.29
Heart rate	1.28	0.33

‡ $p < 0.01$.
† $p < 0.001$.

BSQ scores for both phases: baseline, $F(4,9) = 8.20$, $p < 0.005$; 5% CO_2 inhalation, $F(4,9) = 5.53$, $p < 0.02$. A similar pattern emerged for the ACQ scores with significant R2 emerging for both baseline, $F(4,9) = 18.18$, $p < 0.001$, and 5% CO_2 inhalation, $F(4,9) = 38.18$, $p < 0.001$. Chest EMG accounted for significant variance for BSQ and ACQ scores ($p < 0.05$) during both phases. Minute ventilation accounted for significant variance associated with the BSQ and ACQ scores during baseline ($p < 0.05$) but not during 5% CO_2 inhalation. Heart rate contributed significantly ($p < 0.05$) to baseline BSQ scores only. The contributions of forehead EMG were nonsignificant.

DISCUSSION

Preliminary data are presented indicating that heightened chest EMG activity recorded from the region of the sternum constitutes a physiologic component of panic episodes and a strong predictor of panic-based somatic and cognitive symptoms. By dividing panic disorder subjects on the basis of anxiety level, it was observed that the heightened chest EMG activity was specific to the panic subjects who experienced moderate to severe anxiety during the course of the study. The EMG activity observed in the chest region was not matched by similar EMG activity in the forehead. Heart rate was variable and not significantly different across the groups. Minute ventilation group differences were nonsignificant. These findings need to be interpreted with considerable caution as the group sample sizes were extremely small.

Group differences in chest EMG occurred across all phases of the experiment, including the initial baseline period. The chest EMG measure was nonresponsive to the experimental procedures as there was no increase from baseline during the CO_2 inhalation phases. We cannot determine, from the available data, whether the heightened chest EMG activity was chronic in nature or specific to the experimental situation. Other researchers (10) have noted high anxiety and physiological arousal in panic disorder patients during baseline periods of provocation paradigms. We suspect that the pulmonary laboratory setting proved to be highly anxiety provoking for the

more severe patients, despite the use of antipanic instructions and reassurances.

The HA and LA panic disorder subjects, although grouped on the basis of their anxiety levels during the experimental situation, also differed significantly on a pre-experimental measure of total phobic behavior and agoraphobic avoidance activity. A hallmark clinical feature of agoraphobia is fear of situations that prevent easy escape. Entering the pulmonary lab and having the physiologic transducers attached certainly made a rapid or unnoticed exit difficult. Thus, the severity of the condition may have prompted the panic-related chest EMG activity observed during the session.

In a review of the sodium lactate infusion literature, Margraf et al. (10) commented that baseline differences in arousal may explain why some subjects panic to lactate infusion and other subjects do not. The reviewers also suggested that the baseline differences between panickers and nonpanickers may reflect acute anticipatory anxiety rather than chronic differences in level, with the lactate pushing more highly aroused subjects across a tolerance threshold. We propose, in a similar fashion, that the HA panic disorder subjects in the present study were initially exhibiting high levels of anticipatory anxiety and panic-related somatic and cognitive symptoms. The accompanying chest EMG activity, within this context, was likely a physiological component of the developing panic attack. Panic subjects who showed no anxiety exhibited low levels of chest EMG activity and were very similar to the controls on this measure.

Group differences in heart rate were nonsignificant. Other investigators (9) have observed a lack of relationship between heart rate and subjective reports of panic. The respiratory measure of minute ventilation also failed, in the present study, to statistically differentiate the panic groups and controls. Minute ventilation was predictive of somatic symptoms and cognitions associated with panic during baseline. However, there was no evidence for differences between the groups in terms of responsiveness to the different levels of CO_2 inhalation. The use of the 5-minute inhalation period may not have been sufficient for all subjects to have reached asymptotic levels of minute ventilation. In addition, the small number of subjects prevents reaching a conclusion that there are no differences in CO_2 responses between panic patients and controls.

Given the diagnostic procedures used, we cannot be certain that the present findings are specific or unique to panic disorder. The subject selection/diagnostic procedure did not rule out the possibility of other Axis I conditions. Such a possibility would not, however, change the basic conclusion that chest wall EMG activity is a component of panic anxiety. The group differences in chest wall EMG activity were discernible in spite of the fact that three of the six high anxious subjects were receiving medications for panic. It is possible that these medications may have attenuated group differences in the other physiological parameters.

Freeman and Nixon (11) proposed that overuse and fatigue of the intercostal muscles is responsible for the chest symptoms experienced by panic sufferers. Surface electrodes placed over the sternum can pick up activity from pectoral and parasternal intercostal muscles. In normal subjects, pectoral muscles are used in respiration only for vigorous inspiratory ef-

forts, such as when there is mechanical impedance to breathing. The parasternal intercostal muscles are used for inspiration in normal quiet breathing, but are silent in expiration (12, 13). The EMG reported in this study was the trough of the time-varying signal and corresponded to end-expiration. The sustained EMG activity seen in expiration implies an inspiratory activity maintained throughout the whole breathing cycle. That is, these muscles are holding the chest, in a tonic fashion, at a volume higher than its relaxed volume, plus making repeated inspiratory efforts to increase volume further for each breath. Such activity would be quite abnormal and serve no physiological purpose and could potentially be quite tiring.

Biologic models of panic are viewed by an increasing number of investigators as too narrow and in need of expansion to include both psychological and physiological factors. It has been suggested that biologic challenge tests, such as CO_2 inhalation, rather than directly activating biologic panic sites, operate by eliciting somatic sensations to which patients respond with fear (2); i.e., carbon dioxide challenge leading to bodily sensations that are perceived by panic sufferers as frightening. Furthermore, this viewpoint implies that such challenge tests produce similar physiological responses in everyone and it is the hypervigilance for accompanying somatic sensations which is the significant determinant of panic. The present findings indicate that there may be a musculoskeletal substrate for some of the sensations. Replication and elucidation of the chest wall EMG findings, under conditions of threat and non-threat, would add to our understanding of this condition.

SUMMARY

As a pilot investigation, the level of EMG activity was measured from the sternum of panic disorder patients and controls who were participating in a larger ventilatory study. All subjects were exposed to three concentrations of carbon dioxide inhalation during a single session. Chest muscle activity was recorded from surface electrodes throughout the session. Additional physiological measures recorded during the session included forehead EMG, heart rate, and minute ventilation. Psychological data were obtained prior to the laboratory session and across the experimental phases within the session.

Prior to conducting the main statistical analyses, the panic disorder patients were sub-grouped into High Anxious (HA) and Low Anxious (LA) panic disorder groups on the basis of their average anxiety scores reported during the session. HA panic patients ($n = 6$) had significantly higher chest EMG activity than LA panic patients ($n = 3$) and controls ($n = 8$) across all phases of the study, including baseline. The groups showed no differences in terms of frontalis EMG activity. Group differences on the heart rate and minute ventilation measures were also nonsignificant. A series of regression analyses revealed that chest EMG activity proved to be the strongest physiologic predictor of somatic and cognitive symptoms measured during baseline and in response to 5% CO_2 provocation.

The findings, although very preliminary, were interpreted as support for chest EMG activity as a physiologic component of the panic episode. The activity is believed to be partially independent of the respiratory cycle and may have parasternal and pectoralis muscle origins.

REFERENCES

1. Gorman JM, Liebowitz MR, Fyer AJ, Stein J: A neuroanatomical hypothesis for panic disorder. Am J Psychiatry 146:148-161, 1989
2. Barlow DH: Anxiety and its disorders: The nature and treatment of anxiety and panic. New York, Guilford, 1988
3. Ottaviani R, Beck AT: Cognitive aspects of panic disorders. J Anxiety Dis 1:15–28, 1987
4. Bakal DA, Lynch PB, Whitelaw WA: Carbon dioxide recovery in panic disorder patients and controls. Paper presented at the Annual meeting of the American Psychosomatic Society, Boston, 1990
5. Katon W, Hall ML, Russo J, Cormier L, Hollifield M, Vitaliano PP, Beitman BD: Chest pain: Relationship of psychiatric illness to coronary arteriographic results. Am J Med 84:1–9, 1988
6. American Psychiatric Association: Diagnostic and Statistical Manual, Third Edition. Washington, DC, American Psychiatric Association, 1980
7. Chambless DL, Caputo GC, Bright P, Gallagher R: The assessment of fear of fear in agoraphobics: The body sensations questionnaire and the agoraphobic cognitions questionnaire. J Consult Clin Psychol 52:1090–1097, 1984
8. Marks IM, Mathews AM: Brief standard self-rating scale for phobic patients. Behav Res Ther 17:263–267, 1979
9. Margraf J, Taylor B, Ehlers A, Roth WT, Agras WS: Panic attacks in the natural environment. J Nerv Ment Dis 175:558–565, 1987
10. Margraf J, Ehlers A, Roth WT: Sodium lactate infusion and panic attacks: A review and critique. Psychosom Med 48:23–51, 1986
11. Freeman LJ, Nixon PGF: Chest pain and the hyperventilation syndrome—Some aetiological considerations. Postgrad Med J 61:957–961, 1985
12. DeTroyer A, Sampson MC: Activation of parasternal intercostals during breathing efforts in human subjects. J Appl Physiol 52:524–529, 1982
13. Taylor A: The contribution of the intercostal muscles to the effort of respiration in man. J Physiol (Lond) 151:390–402, 1960

British Journal of Obstetrics and Gynaecology
January 1991, Vol. 98, pp. 84–91

A randomized controlled trial of early amniotomy

WILLIAM D. FRASER, REG SAUVE, I. JOHN PARBOOSINGH, TAK FUNG, ROBERT SOKOL, DAVID PERSAUD

Abstract

Objective—To determine if a policy of early amniotomy resulted in a reduction in mean labour duration when compared to a policy of conservation of the membranes.
Design—A single-centre randomized controlled trial.
Setting—A tertiary care teaching hospital in Alberta, Canada.
Subjects—Ninety-seven term nulliparae in spontaneous labour, baby in cephalic presentation.
Intervention——Early amniotomy versus intent to keep membranes intact.
Main outcome measures—Interval from randomization to delivery, rate of abnormalities of fetal heart rate tracings, cord artery blood pH, Apgar scores.
Results—The mean interval from randomization to delivery was 390·9 (SE 29·1) min in the amniotomy group and 442·9 (SE 34·1) min in the control group ($P = 0{\cdot}251$). There were no differences between groups in the occurrence of fetal heart rate tracing abnormalities, nor was there a difference in the proportion of babies with abnormal Apgar scores, or abnormal cord pH ($< 7{\cdot}20$).
Conclusion—The results of the study fail to support the long held belief that early amniotomy is an effective method for reducing labour duration.

Amniotomy was introduced to obstetric practice by Kreis (1928) who, on the basis of data from case series, argued that it was an effective method to prevent prolonged labour. Amniotomy has recently been promoted as a component of the active management of labour, a protocol designed to reduce both the duration of labour and the occurrence of dystocia (O'Driscoll *et al.* 1984). The effects of amniotomy on the duration of labour have not been adequately assessed. The primary objective of this randomized trial of labour management was to determine the effectiveness of early amniotomy in reducing the duration of labour. The effects of the policy of membrane management on indicators of fetal and neonatal status were also assessed.

University of Calgary, Department of Obstetrics and Gynaecology and Community Health Sciences
W. D. FRASER
I. J. PARBOOSINGH
D. PERSAUD
Community Health Sciences and Pediatrics
R. SAUVE

Academic Computing Services
T. FUNG

Wayne State University
Department of Obstetrics and Gynecology
R. SOKOL

Correspondence: Dr William Fraser, Périnatologie, Hôpital St-François d'Assise, 10 rue de l'Espinay, Québec, QC Canada G1L 3L5.

Subjects and methods

Subjects were eligible if they met the following criteria: nulliparous, spontaneous labour, single fetus in cephalic presentation, term pregnancy (≥ 38 weeks), intact membranes, and a normal fetal heart monitor tracing at admission. Exclusion criteria were a history of genital herpes, the presence of proteinuria or hypertension, or a cervical dilatation of ≥ 5 cm at admission. Demographic information on eligible non-participants was obtained from the hospital's computerized data base.

The study was approved by the University

Medical Ethics Committee and consent was obtained at the time of admission. A non-stratified randomization technique was employed using the block method of Zelen (1974). A series of numbered, sealed, opaque envelopes was prepared containing allocation. Allocation was later verified against the master list. Women with cervical dilatation of < 3 cm were randomized when the fetal head was fixed in the pelvis and the cervix had undergone a change in dilatation after admission. Women with a cervical dilatation of ≥ 3 cm were randomized when the fetal head was fixed in the pelvis.

Women in the amniotomy group underwent the procedure as soon as possible while those in the control group had membranes conserved unless intervention was dictated by the judgment of the treating physician. Amniotomy during labour was permissible in the control group if there was a fetal heart rate abnormality, if there had been an arrest of cervical dilatation for ≥ 2 h, or when full dilatation was achieved. Physicians were requested not to augment labour with oxytocin unless there had been an arrest of cervical dilatation for ≥ 2 h. The frequency and duration of post-randomization electronic fetal monitoring was left to the treating physician. The time from randomization to delivery was the outcome of interest.

Fetal heart rate tracings were later analyzed by two trained observers who worked independently and who were blinded to allocation. The following definitions were used as guidelines in the interpretation of the tracings. Bradycardia: < 110 bpm; tachycardia: > 160 bpm; decreased variability: ≤ 5 bpm about baseline. Decelerations were classified as (1) early≥ 10 bpm below the baseline rate, onset within 20 s of the onset of contraction, recovery to baseline rate within 20 s of the end of the contraction; (2) late: repeated decreases in fetal heart rate, onset > 20 s after onset of contractions, recovery to baseline rate > 20 s after end of contraction; (3) variable: sharp decrease in rate from baseline and a variable relation to contractions. These were subclassified as (i) severe variable—one of the following criteria had to be met: ≥ 60 s duration; ≥ 60 bmp below baseline; a rate of < 60 bmp at its lowest point; (ii) mild variable: ≥ 15 bpm below the baseline and ≥ 15 s duration but failing to meet the criteria for severe variable decelerations. Variable-type decelerations which were < 15 s duration were disregarded. Repeated variable decelerations were counted as separate events only if the fetal heart rate returned to the baseline between decelerations. A dichotomous outcome (normal versus abnormal) was assigned to each hour during which at least 10 min of interpretable tracing was obtained. The segment of tracing was considered 'abnormal' if any of the following features were noted: decreased variability, severe variable or late decelerations, baseline tachycardia or bradycardia. The proportion of 'abnormal' hours was determined for both the pre-randomization and post-randomization periods. For each hour of tracing, the two observers were considered to agree on interpretation if the difference between observers in the total number of decelerations counted for that hour was no more than one, and if there was complete agreement on the normality or abnormality of the baseline fetal heart rate and of its variability. There was agreement between the observers for 87·5% of the hours assessed. When there were disagreements, the tracings were reviewed by the two observers together and a consensus on the interpretation of the tracings was achieved for all hours.

Umbilical artery blood was collected in a heparinized syringe by the delivery physician and was analyzed on a Corning Blood Gas Analyzer within 30 min of delivery. The Apgar score was assigned by the delivery room nurse or by a member of the paediatric staff, if one was present at delivery.

The comparisons of group means were performed using Student's *t*-test. The χ^2 test was used for comparison of proportions, except where the expected value for a cell was less than 5. In this case Fisher's exact test was used. The multiple linear regression analyses used Minitab (Ryan 1985). BMDP-2L (Hoplans 1985) was used to apply Cox's proportional hazards model to the survival data. The power analysis was performed using the statistical package Power (1985).

Results

The study was conducted at Foothills Hospital, a tertiary care teaching centre associated with the University of Calgary, between Sept 11, 1987 and Sept 11, 1988. Of the 1503 nulliparous women delivered during the study period, 278 (18%) were eligible to participate. The reasons for ineligibility and their frequencies are shown in Table 1. Of the 278 women eligible to partici-

Table 1. Reasons for exclusions in 1503 nulliparous women delivered during the study period

	N	(%)
Total no. of women delivered	1503	(100)
Total no. excluded		
Reasons for exclusion:	1225	(81·5)
Spontaneous rupture of membranes	373	(24·8)
Induction of labour	335	(22·3)
Cervical dilation > 5 cm at admission	113	(7·5)
Malpresentation	67	(4·5)
< 266 days gestation	157	(10·4)
Suspected fetal distress	52	(3·5)
Other (including med. complication)	128	(8·5)
Total no. eligible	278	(18·5)

pate 97 (35%) entered the study. The reasons for non-participation of eligible women were patient refusal (44%), doctor refusal (5%) and failure to offer participation by the nursing staff (16%). The mean gestational age at delivery was 39·9% (SD 0·1) weeks for non-participants and 40·0 (SD 0·1) weeks for participants. The mean maternal age was 27·5 (SD 0·3) years for non-participants and 26·7 (SD 0·6) years for participants. Oxytocin use for labour augmentation was similar in the two groups (non-participants 27·6%, participants 30·9%), as was the frequency distribution of the methods of delivery (data for non-participants not shown). A total of 100 women were randomized, 49 to the amniotomy group and 51 to the control group. Eligible patients did not change treatment groups after randomization, regardless of the evolution of their membrane status. Three women who did not meet the eligibility criteria were randomized in error (1 breech, 2 inductions of labour). Of the 97 eligible women who were randomized, 47 were allocated to the amniotomy group and 50 to the control group (conservation of the membranes). As seen in Table 2, the mean prepregnancy weight, the mean admission weight and the mean birthweight were somewhat greater in the amniotomy group than in the control group. The groups were similar with respect to maternal height education, and gestational age. The mean cervical dilatation at admission was somewhat less in the amniotomy group than in the control group (Table 2). A similar proportion of women in the two groups smoked during pregnancy, attended prenatal classes, was accompanied by a significant other (usually the husband), and claimed English as their mother tongue (data not shown).

Of the 47 women in the amniotomy group two had a spontaneous rupture of the membrane before amniotomy. Of the 50 women in the control group 19 had an amniotomy before full dilatation. The indications for amniotomy in the control group were labour augmentation (11), and suspected fetal distress (8). The mean interval from labour onset to rupture of the membranes was 431 (SE 36) min in the amniotomy group and 592 (SE 37) min in the control group ($t_{95df} = 3·13$, $P = 0·002$). Figure 1 is an histogram of the distribution of cervical dilatations at the time of membrane rupture for the two groups. Of the women in the control group 60% achieved ≥ 8 cm cervical dilatation before membrane rupture compared with only 2% in the amniotomy group.

Fifteen women in each group had labour augmented with oxytocin. There were eight caesarean sections in the amniotomy group and four in the control group. Similar proportions of women in each group were delivered by low forceps or vacuum extractor (amniotomy 23%, control 28%), and by mid-forceps (8% in each group).

The mean interval from randomization to

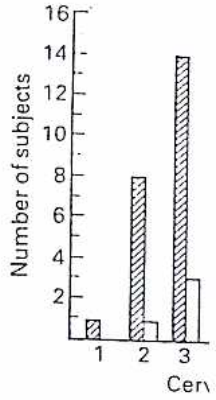

Fig. 1. Frequency dist
the time of rupture
(shaded) and control

delivery was 390·9
tomy group and 44
trol group (t_{95df} =
analysis of labour d
group of women wh
mean duration of la
labour for these wo
ment group in Tabl
domization to full
significantly betwee

The greater mean
at admission and
mean cervical dila
amniotomy group
longer labours in
between these varia
randomization to
delivered vaginall
multiple linear regre
sion coefficients that
able were obtained
the last one in the
treatment policy of
not a statistically s
duration of the inter

To better describ
treatment groups w

Table 3. Duration of la

Onset-delivery (min)
Onset-admission (min)
Admission-randomizat
Randomization to full
Second stage (min)

Table 2. Maternal demographic data for study patients: amniotomy and control groups

	Amniotomy (n = 47)		Control (n = 50)	
	Mean	(SE)	Mean	(SE)
Years of education	13·3	(0·3)	14·3	(0·4)
Prepregnancy weight (kg)	61·1	(1·75)	57·3	(7·1)
Admission weight (kg)	76·3	(1·7)	71·8	(1·0)
Height (cm)	164·8	(0·9)	165·3	(0·9)
Birthweight (g)	3493·1	(73·0)	3322·1	(64·0)
Gestational age (days)	281·1	(0·8)	279·5	(1·0)
Cervical dilatation first exam (cm)	2·4	(0·02)	2·8	(0·02)
Onset of labour to randomization (min)	399·0	(31·0)	331·0	(28·0)

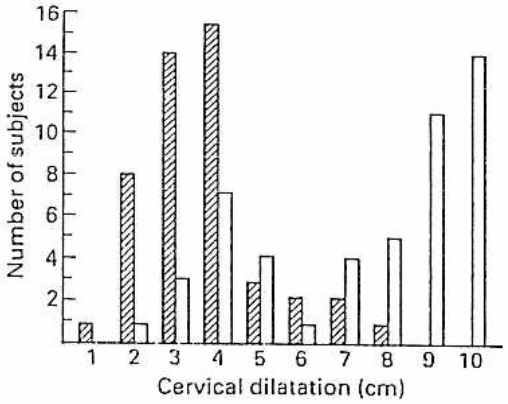

Fig. 1. Frequency distribution of cervical dilatation at the time of rupture of membranes in amniotomy (shaded) and control (unshaded) groups.

delivery was 390·9 (SE 29·1) min in the amniotomy group and 442·9 (SE 34·1) min in the control group ($t_{95df} = 1·15$, $P = 0·251$). A further analysis of labour duration was based on a subgroup of women who had vaginal deliveries. The mean duration of labour and the sub-intervals of labour for these women are displayed by treatment group in Table 3. The interval from randomization to full dilatation did not differ significantly between the two groups.

The greater mean values for maternal weight at admission and birthweight and the lower mean cervical dilatation at admission in the amniotomy group could have biased toward longer labours in that group. The relations between these variables and the interval from randomization to delivery in the women delivered vaginally were investigated by multiple linear regression (Table 4). The regression coefficients that are expressed for each variable were obtained by entering that variable as the last one in the regression equation. The treatment policy of membrane management was not a statistically significant predictor of the duration of the interval.

To better describe the behaviour of the two treatment groups with respect to the interval from randomization to delivery, survival curves were created. For the purposes of this analysis, the critical event was defined as vaginal delivery. Women delivered by caesarean section were considered as censored following the procedure. Cox's proportional hazards model was applied to the data to control for differences in co-variates that could influence labour duration. The variables included in the model along with their regression coefficients are shown in Table 5. A positive coefficient indicates an increase in the hazard function and an inverse relation with survival (continuation in labour). The graphic representations of the adjusted estimates of the survival functions are shown in Fig 2. Trial status (amniotomy versus conservation of the membranes) was not a statistically significant predictor of survival function, adjusting for birthweight, cervical dilatation at admission and maternal admission weight.

With respect to maternal post-partum morbidity, no women required blood transfusion. One woman in the amniotomy group developed a wound cellulitis after caesarean section and was treated with oral antibiotics. No woman met the study criteria for febrile morbidity (two episodes of fever > 38° C on two occasions, excluding the first 24 h after delivery).

Fifty five women in the amniotomy group and 47 women in the control group had post-randomization fetal heart tracings. The mean duration of electronic fetal monitoring was similar in the two treatment groups (amniotomy: 7·73 SE 0·47 h; control: 8·57, SE 0·48 h). Concerning the dichotomous fetal heart rate tracing outcome, the mean proportion of hours during which the tracings were 'abnormal' was similar in the two groups (post-randomization: amniotomy: mean 0·41, SE 0·05; control mean 0.44, SE 0.04; $t = 0·50$, $P = 0·620$).

The two treatment groups were similar with respect to the frequency of early, mild variable, severe variable and late decelaration during the pre-randomization period. The mean hourly

Table 3. Duration of labour (minutes) in women delivered vaginally in the amniotomy and control groups

	Amniotomy ($n = 39$)		Control ($n = 46$)				
	Mean	(SE)	Mean	(SE)	t	df	P
Onset-delivery (min)	776	(52)	763	(48)	0·21	83	0·842
Onset-admission (min)	201	(26)	192	(31)	0·21	83	0·842
Admission randomization (min)	192	(29)	149	(16)	1·37	83	0·175
Randomization to full dilatation (min)	311	(20)	346	(30)	0·82	83	0·422
Second stage (min)	72	(9)	75	(6)	0·24	83	0·807

Table 4. Multiple linear regression examining several variables for effect on interval from randomization to delivery in the women delivered vaginally

Variable	Regression coefficient	*t*-ratio	*P*
Constant	264·7		
X1			
Birthweight	0·14	2·88	0·005*
X2			
Cervical dilatation at admission	−39·10	−1·71	0·091
X3			
Maternal admission weight	−2·64	−0·98	0·332
X4			
Trial status (0 = control, 1 = amniotomy)	−80·10	−1·67	0·099

Interval (randomization to delivery) = 264·7 + 0·14(X1) − 39·1 (X2) − 2·64 (X3) − 80·10 (X4)

*Significant at the alpha = 0·01 level

rates of the four types of deceleration during the post-randomization period are shown in Table 6. There were no statistically significant differences between the two groups in the rates of deceleration after randomization.

The results of the variables which are indicators for condition at birth are summarized in Table 7. The proportions of babies with a 1-min Apgar score of < 6 and a 5-min Apgar score of < 8 did not differ significantly between the two groups. Umbilical artery blood pH was measured in 42 of the 47 babies born in the amniotomy group and in 47 of the 50 babies born in the control group. No statistically significant difference was found between the two groups in the proportion of infants with pH < 7·20. The mean cord artery pH was 7·254 (SE 0·07) in the amniotomy group and 7·256 (SE 0·08) in the control group. Six infants in each group were provided with assisted ventilation at birth. Neonatal cephalohaematoma was considered present if the diagnosis was recorded in the baby's chart by the treating physician. This occurred in one infant in the control group who was born by mid-forceps after a prolonged second stage. There was no abnormality noted on skull x-ray. This infant was cared for in the normal nursery and was discharged from hospital without further complications. All babies left hospital alive and in good condition.

Discussion

Several controlled studies which assess the effect of amniotomy on labour duration have been reported. Schwarcz & Caldeyro-Barcia (1982) reported a multi-centre study of 1124 nulliparous and multiparous women. The mean duration of the active phase of labour (5 cm to full dilatation) was 165 min in women who had an amniotomy compared with 213 min in a control group ($P < 0·001$, effect size = 0·77). This study suffers from methodological problems as recently reviewed by Keirse (1989). A smaller randomized trial by Stewart *et al.* (1982) found mean labour duration (admission to full dilatation) to be 4·9 h in the experimental group and 7·0 h in the control group ($P < 0·02$, effect

Table 5. Analysis using Cox's proportional hazards model to assess the effects of variables on survival function

Variable	Coefficient	Coefficient/SE	*P*
X1			
Birthweight	−0·001	−2·476	0·013*
X2			
Cervical dilatation at admission	0·118	1·158	0·2468
X3			
Maternal admission weight	0·012	1·005	0·3188
X4			
Trial status (0 = control,	0·231	0·960	0·337

* Significant at the alpha = 0·05 level.

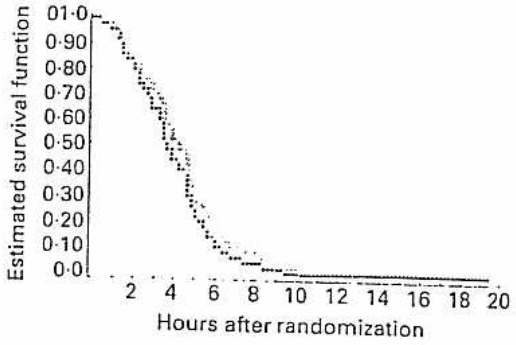

Fig. 2. Estimated probability of continuation in labour for each hour following randomization using Cox proportional hazards model (Amniotomy—closed circles, control—open circles.)

size = 0·7). The authors did not stratify for parity in the analysis. Two smaller studies by Wetrich (1970) and Laros (1972) reported similar results.

In that our study was designed to assess the effectiveness of a commonly used obstetrical procedure in a low risk population, it was surprising that the proportion of women who were eligible to participate was so small. The high number of exclusions was, in part, a function of the study having been performed in a tertiary care centre. There was a significant number of refusals by eligible patients, the reasons for which were not assessed. Because recruitment was completed in early labour, it was considered important that women be informed that they were free either to participate or not to participate. Our ability to compare study participants with eligible non-participants is limited by the few variables which were measured in the non-participant population. However, the available data suggest the two groups were similar.

In the course of the main analysis, differences between the treatment groups were identified on several co-variates. This is a cause for caution in the interpretion of the results. The comparisons regarding labour duration which are first presented are based on the unadjusted data. In general, the adjusted analyses served to support the conclusions which had been made on the basis of the unadjusted comparisons. With respect to the duration of labour, comparison of the unadjusted as well as the adjusted means provides no support for rejection of the null hypothesis. The analyses based on the survival curves further strengthen these conclusions.

The initial sample size calculation was based on the formula $n = 2(Z\alpha + Z\beta)^2/(\log_e E)^2$ where E is the effect size (Donnor, 1984). E = 0·7 was used for this calculation. This formula, which is designed for survival analyses and does not take into account the dispersion of the data, yielded a sample size requirement of 110 subjects/group. Because of limited resources, a power analysis was performed after 97 subjects had been recruited using a formula which takes into account both the dispersion of the data and the group means that had been attained. The calculation indicated that a power of 0·85 had been achieved for the specified effect size (0·7). It was decided to terminate the study at that point. It should be noted, however, that the power of the study to detect a 20% reduction in labour duration was only 0·55.

Many of the women in the current study were randomized when cervical dilatation was < 3 cm. This may partially explain the discrepancy between the findings of the current study and that of previous clinical trials when women were randomized later in labour. Further studies of this type should stratify by cervical dilatation at the time of randomization to assess whether any observed effect on labour duration is dependent on the cervical dilatation at the time of the procedure.

Several previous reports have suggested an association between early amniotomy and the occurrence of fetal heart rate decelerations. Schwarcz *et al.* (1973) found that after amniotomy over 20% of contractions produced type I

Table 6. Mean hourly rates of early, mild variable, severe variable, and late decelerations during the post randomization period for patients in the amniotomy and control groups

	Amniotomy		Control				
Type of deceleration	Mean	(SE)	Mean	(SE)	t	df	P
Early	0·41	(0·17)	0·49	(0·23)	0·32	90	0·750
Mild variable	1·10	(0·20)	1·50	(0·32)	1·23	90	0·220
Severe variable	1·61	(0·33)	2·83	(0·61)	1·73	90	0·084
Late	0·32	(0·22)	0·54	(0·40)	0·47	90	0·640

Table 7. Indicators of early neonatal status for infants born in the two study groups

	Amniotomy		Control		Fisher's exact
	Total no.	No. affected (%)	Total no.	No. affected (%)	P-value
Apgar 1 min < 6	47	7 (14·9)	50	9 (18·0)	0·787
Apgar 5 min < 8	47	2 (4·3)	50	5 (10·0)	0·437
Cord artery blood pH < 7·20	42	8 (17·0)	47	6 (12·0)	0·566
Assisted ventilation	47	6 (12·8)	50	6 (12·0)	1·000
Admitted to intensive care nursery	47	1 (2·1)	50	2 (4·0)	1·000

(early) decelerations, as opposed to 3% of contractions before membrane rupture.

Baumgarten (1976) reported the occurrence of early decelerations to be approximately doubled in women undergoing amniotomy relative to those labouring with membranes intact. Aladjem & Miller (1977) in an observational study found that spontaneous or artificial rupture of the membranes produced transient FHR abnormalities in approximately 25%, mainly early decelerations. By 15 min after rupture of membranes, there were no differences between groups in the occurrence of FHR abnormalities. Stewart *et al.* (1982), on the other hand, failed to show any differences between treatment groups in the rate of FHR tracing abnormalities, although the methods used to classify the tracings were not clearly specified.

With respect to the effects of early amniotomy on neonatal outcome, Schwarcz *et al.* (1982) found no association between the state of the membranes and Apgar score or neurological condition at 48 h of life. Baumgarten (1976) and Stewart *et al.* (1982) found no differences between groups in Apgar scores or in cord artery blood pH. Martell *et al.* (1976), in a study of 38 patients, reported early membrane rupture to be associated with a lower mean cord artery blood pH than occurred in infants of patients who laboured with intact membranes.

There are several advantages of the current study over previous studies with respect to indicators of fetal and neonatal status. The randomization process was carefully blinded, and the methods of classification of the FHR tracings was clearly specified. The dichotomous FHR tracing outcome selected as abnormal only those findings which could be of clinical significance to obstetrical decision making. As is appropriate in the analysis of management trials, the participants remained in their respective groups despite the post-randomization evolution of membrane status. Although this approach increases the risk of beta error, the absence of trends in the data suggests that it is unlikely that larger studies would demonstrate differences between treatment groups in fetal status or in neonatal outcomes.

In conclusion, the results of this study fail to support the longheld belief that early amniotomy is an effective method to reduce labour duration. A multicentre trial designed to further assess the effects of the procedure on labour progress is underway.

Acknowledgment

During the period of this study, Dr Fraser was supported by the Alberta Heritage Foundation for Medical Research.

References

Aladjem S. & Miller T. (1977) Effects of spontaneous and artificial membrane rupture in labour upon fetal heart rate. *Bre J Obstet Gynaecol* 84, 44-47.

Baumgarten D. (1976) Advantages and disadvantages of low amniotomy. *J Perinatol Med* 4, 3-9.

Caldeyro-Barcia R. *et al.* (1974) Adverse perinatal effects of amniotomy during labour. In *Modern Perinatal Medicine*, (Gluck L. ed), Yearbook Medical Publishers, Chicago, pp. 431-449.

Donner A. (1984) Approaches to sample size estimation in the design of clinical trials—A review. *Statistics in Medicine* 3, 199-214.

Hoplans S. (1985) Survival analysis with covariates. cox model. In *MBDP Statistical Software Manual. (Dixon W. J., Brown M. B., Englemeni L., et al.)*. University of California Press, Berkeley, pp. 557-75.

Keirse M. J. N. C. (1989) Augmentation of labour. In *Effective Care in Pregnancy and Childbirth* (Chalmers I., Enkin M., Keirse, M. J. N. C. eds), Oxford University Press, Oxford, pp. 950-966.

Kreis J. (1928) L'accouchement médical. *Rev Fr Gynecol Obstet* 19, 604-616.

Laros R. K., Work B. & Witting W. C. (1972) Amnio-

tomy during the active phase of Labour. *Obstet Gynecol* 39, 702-704.

Martell M., Belizan J. M. *et al.* (1976) Blood acid base balance at birth in neonates from labors with early and late rupture of membranes. *J Pediatr* 89, 963-967.

O'Driscoll K., Foley M. & MacDonald D. (1984) Active management of labor as an alternative to cesarean section for dystocia. *Obstet Gynecol* 63, 485-490.

Power (1985) *Epicentre Software*, Pasadena California.

Ryan B. F., Joiner G. L. & Ryan T. A. (1985) *Minitab Handbook* 2nd Ed. Duxbourg Press, Boston.

Schwarcz R. & Caldeyro-Barcia R. (1982) *Amniotomy*. Scientific publication CLAP 931. Latin American Centre of Perinatology and Human Development. Pan American Health Organization. World Health Organization.

Schwarcz R., Althabe O., Belitzky R. *et al.* (1973) Fetal heart rate patterns in labours with intact and with ruptured membranes. *J Perinat Med* 1, 153-165.

Stewart P., Kennedy J. H. & Calder A. A. (1982) Spontaneous labour: When should the membranes be ruptures? *Br J Obstet Gynaecol* 89, 39-43.

Wetrich D. (1970) Effect of amniotomy upon labour. *Obstet Gynecol* 35, 800-806.

Zelen M. (1974) The randomization and stratification of patients to clinical trials. *J Chronic Dis* 27, 365-375.

Received 9 October 1989
Accepted 4 July 1990

Predictors of Community Discharge from a Geriatric Assessment and Rehabilitation Unit

L. Pereles, J. Peitchinis, E. Jennings and T. Fung, *Holy Cross Hospital*

RÉSUMÉ

Cette étude examine les facteurs d'admission pouvant déterminer la réinsertion dans la communauté de patients ayant été admis au programme gériatrique d'évaluation et de rééducation (GARP). La régression logistique à inclusion hiérarchique a été utilisée afin de déterminer les facteurs de prédiction au sein d'un échantillon consécutif de 100 personnes, nécessitant des soins médicaux ou chirurgicaux en gériatrie, admises à un programme de rééducation de quatre à six semaines. Les principaux prédicteurs indépendants établissant le congé et pouvant être déterminés dès l'admission étaient: la possibilité de prescrire sans danger des médicaments, l'admission d'une personne provenant de la communauté, la valeur «GDS» et le soutien social. La possibilité de prescrire sans danger des médicaments peut constituer un prédicteur important, quoi qu'il ne soit pas souvent utilisé.

ABSTRACT

This study explored admission factors which predicted the successful return to the community of patients entering a Geriatric Assessment and Rehabilitation Program (GARP). A stepwise logistic regression technique was used to determine predictive factors from a consecutive sample of 100 medical or surgical geriatric patients admitted for a four to six week rehabilitation program. The significant independent predictors of discharge which could be determined at the time of admission were ability to safely medicate, admission from the community, GDS score and the number of supports. The ability to safely medicate may be an important but under-utilized predictor.

Introduction

Geriatric Assessment and Rehabilitation units are rapidly profilerating in

Key Words: Community Discharge, Geriatric Assessment Rehabilitation.
Mots clés: Congé, évaluation gériatrique de rééducation.
Manuscript received July 8, 1992; manuscript reçu le 8 juillet, 1992.
Manuscript accepted March 11, 1993; manuscrit accepté le 11 mars, 1993.
Requests for reprints should be sent to:/Les demandes de reproduction doivent être adressées à:
L. Pereles
Department of Research
Calgary District Hospital Group
Holy Cross Hospital
2210 2nd Street S.W.
Calgary, AB T2S 1S6

many hospital settings.[1] They have been shown to improve patients' physical and psychological functioning,[2,3,4,5,6] and in some cases to prevent hospital readmission and to reduce mortality.[2,6] These units, however, are fairly costly to operate and access is often limited.[7]

Several authors have suggested that depending on the goals and objective of the program certain types of patients may benefit more than others and that the resources of these units would be better utilized if patients could be more appropriately matched.[6,8,9,10] Previous research has indicated that mental status, functional status and number of supports are important determinants for patients returning to the community.[4,5] This paper will explore these factors and others that could potentially predict which patients are likely to return to the community.

Methods

An evaluation study of the Geriatric Assessment and Rehabilitation Program (GARP) at the Colonel Belcher Hospital in Calgary was conducted from September 1989 to October 1990. One of the purposes of the evaluation was to determine how the patients' physical, psychological and functional status on admission predicted their re-entry into the community.

GARP

The Geriatric Assessment and Rehabilitation Program was established in 1987. It admits medical or surgical geriatric patients with recent deterioration in their functional ability (usually within the last three months) who would benefit from interdisciplinary assessment and a four to six week rehabilitation program and who otherwise would be a candidate for a long-term care facility. The unit excludes patients with acute illness, major psychiatric illness or long-standing dementia. Returning elderly patients to the community is a primary goal of this program.

The primary multidisciplinary team consists of a geriatrician, family physicians, nursing staff, physiotherapist, occupational therapists, recreational therapists, social worker, dietician and pharmacist. Consultants from medical and surgical sub-specialities as well as psychiatry, neuropsychology, and speech therapy are readily available.

Upon admission to the unit the patient is assessed by the team members who prepare a comprehensive care plan. At weekly meetings, the team discusses the patient's progress in individual and group treatments and plans his/her discharge. The patient and his/her family meet with the care team just prior to discharge. The average length of stay is 40 days. Median waiting period is 20 days.

To measure treatment effect, a trained research assistant administered the Katz Activities of Daily Living scale (ADL),[11] Geriatric Depression Scale (GDS),[12] Folstein Mini-Mental State Examination (MMSE),[13] within 48 hours of admission, at discharge, and at a three month post discharge home visit. Mobility was assessed on admission at three levels: independent; im-

paired, requiring the assistance of an aid or person for ambulation but able to transfer independently; impaired ambulation and requiring personal assistance for transfers.

The ability to safely medicate was determined by a pharmacy assessment conducted in the hospital. The patient was assessed as safe if there was a caregiver to administer the medications or if the patient was able to demonstrate on a pharmacy assessment the ability to self-medicate. The pharmacy assessment is standardized and validated assessment which assess a patient's knowledge, memory and functional abilities related to medication administration.[14] If there were any question of the patient's ability to self-medicate, the patient was assigned to a three week self-medication program in which the patient, monitored by the nursing staff, assumes increasing responsibility for the administration of his/her own medication.

Return to the community was defined as discharge to either the patient's own home or family home, senior apartment or lodge and remaining there for three months.

The following demographic information was obtained: age, sex, marital status, number and type of supports (patient indicated who, in the community, helped them and what kinds of help they received). Admitting source (community or post acute care), waiting period prior to admission, and number of medications were also noted.

Sample

One hundred and forty-eight consecutive patient admissions over a one year period were approached for the study. One hundred and twenty-seven were entered into the study. Patients excluded were either untestable, too ill to be tested within the 48 hour admission period or lived outside the study area and could not be tested at three months. Of the 127, eight refused follow-up after entering the study and 19 did not complete the program. There were three cases excluded due to missing data (see Table 1).

Statistical Methods

A backward stepwise logistic regression technique,[15] was used to determine which factors predicted discharge to the community.

The variables which were selected have been previously reported as being predictors in the literature or have been found to be useful in our clinical experience. These were sex; age; marital status; number of supports; waiting period prior to admission; number of medications; the ability to safely self-medicate; mobility; admitting source; admission Katz, MMSE and GDS scores.

Correlations between variables were determined using the Pearson correlation for parametric data and Spearman's rank correlation for nonparametric data. Data was analysed using BMDP statistical software.

Table 1
Study population

Total no. of patients	*148*
Reasons for Exclusion	
Unsuitable due to:	
Aphasia	3
Poor English skills	5
Inadequate comprehension	2
Severe illness	2
Previous inclusion in study	1
Lived outside of study area	8
Dropped out* due to:	
Medical complications	12
Non-completion of program	4
Researcher missing discharge	3
Patient refusal	8

*Three cases were omitted from the regression analysis due to missing date.

Table 2
Functional status of patients and no. of medications

Mean and S.E. Variable	*Admission*	*Discharge*	*Follow-up*
KATZ	2.27 ± .18	1.4 ± .17	1.8 ± .18
MMSE	23.5 ± .52	24.3 ± .48	23.9 ± .55
GDS	12.5 ± .69	10.9 ± .69	11.4 ± .65
NO. OF MEDS	4.0 ± .26	4.6 ± .25	4.4 ± .26
Mobility independent	33%		
Mobility partially dependent	33%		
Mobility dependent	34%		
Ability to safely medicate	75%		

Results

The average age was 79 (range 60–90). There were 62 females and 38 males. Twenty-eight per cent were married; 63 per cent were widowed, separated or divorced and 9 per cent were single. Sixty per cent were admitted post acute care and 40 per cent were admitted directly from the community.

In terms of diagnostic groups, 75 per cent were considered primarily medical; 15 per cent post surgical; and 10 per cent psychiatric. Sixty-eight per cent were discharged back to the community, 28 per cent to long-term care facilities and 4 per cent to acute care. Sixty-six per cent were still in the community at three month follow up.

Patients had an average of two supports. Sixteen per cent of their supports provided direct care while 95 per cent gave emotional support. Seventy-five per cent of patients were able to safely medicate.

Upon admission, 33 per cent of patients were independently mobile, 33 per cent were dependent but still able to transfer and 34 per cent were de-

Table 3
Determinants of return to the community of those completing the study

Predictor	*COEF*	*S.E.*	*t*	*p*	*OR*	*95% CI*
1. Admission from the community	1.05	.35	2.99	.004	8.2	2.0-11
2. No. of supports	.60	.33	1.82	.072		
3. Admission GDS	-.136	.047	-2.92	.004		
4. Ability to safely medicate	1.15	.35	3.28	.001	9.9	2.5-14
Constant	1.29	.95	1.37	.174		
Goodness fit chi sq (2*O LN (O/E) = 79.07 *d.f.*94 *p* = .865						

pendent and unable to transfer. Patients demonstrated small improvements in functional status and retained these improvements over time (Table 2). Two equations were derived: the first included patients who completed the program and had a three month follow-up visit and the second included all those who entered the study but subsequently dropped out, mainly due to illness or refusal of the follow-up visit.

The most significant predictors of discharge, in those that completed the study, in descending importance are: the ability to safely medicate; admission from the community; GDS score; and the number of supports. There were no significant first order interactions between these variables (Table 3).

Many of the variables were significantly correlated with each other due to the large sample size but the correlation coefficients were low except for mobility and ADL = 0.69 (Table 4). To assess possible multicollinearity of these variables, two separate equations, one including mobility (excluding ADL) and one including ADL (excluding mobility) were performed. Multicollinearity was not a problem.

If all persons including the drop-outs are analysed, the predictors remain the same except that ADL replaced the number of supports as a significant predictor. There were no significant first order interactions (Table 5).

Discussion

The advantages of examining those factors which predict community discharge are two-fold. The identification of actors which prevent discharge can result in new rehabilitation strategies to overcome these obstacles. For example, if safe medication administration is the only impediment to returning to the community, then creative solutions may be developed to deal with this problem.

The second advantage is the selection and priorization of patients. Most programs informally screen for high risk patients. Rubenstein et al.[2] excluded patients who could not perform more than three ADL functions, had severe dementia and lacked a social support system. Applegate et al.[6] also excluded patients with severe mental impairment and inevitable nursing

Table 4
Correlations of predictors of community discharge

	D/C-COMM	SEX	AGE	ADM-COMM	ADM ADL*	MMSE
D/C-COMM	1.00	-.01 ns	-.06 ns	0.39 $p=.001$	.29 $p=.004$	0.27 $p=.007$
AGE	-.06 ns	.25 $p=.01$	1.00	-.13 ns	-.18 ns	-.24 $p=.02$
ADM.-COMM	.39 $p=.000$	.09 ns	-.13 ns	1.0	.34 $p=.001$	.31 $p=.002$
NO. OF SUPPORTS	.25 $p=.01$	.06 ns	.04 ns	.03 ns	.12 ns	.04 ns
ADM ADL*	.29 $p=.004$	.14 ns	-.19 ns	.34 $p=.001$	1.0	.32 $p=.001$
ADM MMSE	.27 $p=.007$	.13 ns	-.24 $p=.02$	.31 $p=.002$	.32 $p=.001$	1.0
ADM GDS	-.30 $p=.002$	-.02 ns	.00 ns	.02 ns	-.02 ns	-.15 ns
WAIT PERIOD	.03 ns	-.05 ns	-.1 ns	.23 $p=.02$	.13 ns	.18 ns
SAFELY MED.	.46 $p=.001$	-.12 ns	-.25 $p=.01$	.28 $p=.001$	.32 $p=.001$	.34 $p=.000$
ADM MEDS	.10 ns	.00 ns	-.16 ns	-.03 ns	.08 ns	.14 ns
MOBILITY	.24 $p=.01$	.11 ns	-.11 ns	.31 $p=.001$	.69 $p=.001$	.24 $p=.02$

	ADM-GDS	LENGTH OF STAY	NO.-MEDS.	SAFELY MED.	MOBILITY
D/C-COMM	.30 $p=.002$	-.37 $p=.001$	.02 ns	.46 $p=.001$	.24 $p=.01$
AGE	.005 ns	-.04 ns	-.23 $p=.02$	-.25 $p=.01$	-.11 ns
ADM.-COMM	.02 ns	-.37 $p=.001$	-.02 ns	.28 $p=.005$	.31 $p=.001$
NO. OF SUPPORTS	-.13 ns	-.08 ns	-.07 ns	.13 ns	.13 ns
ADM ADL*	-.02 ns	-.39 $p=.001$	.00 ns	.32 $p=.001$	.69 $p=.001$
ADM MMSE	-.14 ns	-.06 ns	.15 ns	.34 $p=.004$	.24 $p=.007$
ADM GDS	1.0	.11 ns	.15 ns	-.02 ns	-.03 ns
WAIT PERIOD	.15 ns	-.06 ns	.12 ns	.03 ns	.12 ns
SAFELY MED.	-.02 ns	-.18 ns	.20 $p=.05$	1.0	.24 $p=.02$
ADM MEDS	.14 ns	.02 ns	.15 ns	.34 $p=.001$	.06 ns
MOBILITY	-.03	-.31	-.05	.24	1.0

*ADM ADL = "Independence in ADL".
As the Katz ADL uses an inverse scoring procedure (increase in ADL score = decrease in independence) the term "independence of ADL: is used to keep the direction of the ADL correlations in line with the other measures.

Table 5
Determinants of return of the community of all subjects entered into study including drop-outs

Predictor	*COEF*	*S.E.*	*t*	*p*	*OR*	*95% CI*
1. Admission from the community	1.4	.53	2.66	.009	4.1	2.4-6.9
2. ADL	-0.31	1.35	-2.26	.026		
3. GDS	-0.120	.038	-3.20	.0018		
4. Ability to safely medicate	2.09	.59	3.54	.0006	8.1	4.5-14.6
Constant	0.911	.761	1.2	0.2326		

Goodness fit chi sq (2*O LN (O/E) = 108.9 *d.f.*118 $p = .713$

home placement. They concluded that patients at moderate rather than high risk of nursing home placement had the greatest improvement with respect to mortality and functional status. Rehabilitation programs should target those patients who would best match their goals and resources. Realistic and clear establishment of goals prior to admission is important. The rehabilitation goals for patients returning to the community are different than those clearly destined to nursing home placement. It may be necessary to separate and stratify rehabilitation programs based on patients' discharge prospects.

Many of the characteristics of patients who returned to the community in this study are not surprising. Functional and mental status are well-established predictors.[3,4,10,16] Narain et al.[17] in a prospective multivariate study of elderly patients. found that low functional and mental status, living location, and type of caregiver were most predictive of nursing home admission. Functional status was a stronger predictor of length of stay, mortality and nursing home placement than the principal admitting diagnosis. In the present study, functional status as measured by ADL was only a significant predictor when those patients who did not complete the study, mainly due to illness, were included. Not surprisingly, their mean functional status (ADL) was lower than the group who did (2.7 vs 2.3).

Mental status did not appear as an independent predictor. However, it was correlated with the ability to safely medicate and admission from the community. Mental status has not been shown to consistently improve with admission to a geriatric unit. [2,4,6] The number of cases of reversible cognitive impairment may actually be quite small.[18] Moderate to severe impairments in mental status without any reasonable prospect for reversal may be a reasonable exclusion criterion for programs which aim to return patients to the community.

Admission from the community, rather than post-acute care, strongly predicted return to the community. This finding may suggest the importance of earlier intervention for the geriatric patient before his/her condition has deteriorated to the point of requiring an acute care admission. The collapse of community support systems may also be circumvented by earlier intervention.

A patient who returns home will have to have a reasonable functional level for self care, or, if not, a support system to assist with that care. The number of supports is usually a predetermined factor but should be given consideration at the time of admission. In this study the number of supports predicted eventual return to the community. In general, the lack of a caregiver or spouse increases the risk of institutionalization.[4,19,20]

A factor which may not be recognized and given due consideration is depression.[5] Depression was a relatively independent predictor of patients not returning to the community, being only correlated with Mini-Mental Status scores. The prevalence of depression has been estimated to be 15 per cent in the elderly.[21] Depressed, lonely and isolated elderly patients may not be able to return to the community despite a reasonable functional level. Depression can reduce the motivation to participate in rehabilitation programs as well. Because of the significant delay in the response of depression to treatment and its effect on mental status, early identification and treatment of depression is important.

Instrumental Activities of Daily Living (IADL) may be under-utilized predictors. Wachtel et al.[22] examined IADL as determinants of community discharge. Using the ability to do heavy housework independently or climb stairs as predictive measures, he found that 97 per cent of patients who could perform either of these activities were discharged back to the community. IADL by their very nature often require intact cognitive skills and functional independence. In the community, the performance IADL often will determine a person's capability for independent living. The use of IADL's rather basic functions may simplify testing and more accurately reflect outcomes.

The only IADL function measured in this study was ability to safely medicate. It has not been studied extensively as a predictor of hospital discharge although it is a definite factor in decision-making about discharge. Medication misadventures are a common cause of hospitalization in the elderly.[23,24]

The ability to safely medicate is a two-fold variable. It includes situations where the patient can administer his/her own medication safely him/herself or, if not, has assistance to administer them. This definition was used as it corresponds more closely to the clinical reality. The ability to self-medicate may be one of the few instrumental activities of daily living that can be easily measured in a hospital setting. Programs have now been designed to teach patients to self-medicate. Whether such programs are effective still remains to be demonstrated. Other instrumental activities may also be predictive of discharge but they have not been tested. Driving, shopping, and financial management might be potential predictors.

The present study has several limitations. The generalizability of the findings is unknown. The manpower and programming of various geriatric assessment units is not uniform across the country. The treatment outcomes and discharge to the community may also reflect the strengths and weaknesses of individual teams and the availability of community programs and supports. This was an exploratory study: the predictive value of the model has not been validated in a prospective study. Patients were only fol-

lowed for a three month period. Whether the model is valid for longer periods has not been demonstrated.

Conclusion

The determinants of discharge from geriatric assessment and rehabilitation program to the community are multi-factorial. This study suggests that many may be determined prior to admission, thus improving patient selection and resource allocation. Highlighted in this study were the importance of the ability to self-medicate, psychological state and community supports.

References

1. Epstein AM, Hall JA, Besdine R, Cumella Jr. E, Feldstein M, McNeil B et al. The emergence of geriatric assessment units the "new technology of geriatrics". Annual of Internal Medicine 1987; 106:299–303.
2. Rubenstein LZ, Josephson KR, Wieland GD, English MS, Sayre JA, Kane RL. Effectiveness of a geriatric evaluation unit: A randomized clinical trial. New England Journal of Medicine 1984; 311(26):1664–7.
3. Rubenstein LZ, Wieland D, English P, Josephson, K, Sayre JS, Abrass IB. The Sepulveda VA geriatric evaluation unit: Data on four-year outcomes and predictors of improved outcomes. The Journal of the American Geriatric Society 1984; 32(7):503–12.
4. Liem PH, Chernoff R, Carter WJ. Geriatric rehabilitation unit: A 3-year outcome evaluation. Journal of Gerontology 1986; 41(4):44–50.
5. Applegate WB, Akins D, Zwagg RV, Thoni K, Baker MG. The geriatric rehabilitation and assessment unit in a community hospital. The Journal of the American Geriatric Society 1983; 31(4):206–10.
6. Applegate WB, Miller ST, Graney MJ, Elam, JT, Burns R, Akins DE. A randomized, controlled trial of a geriatric assessment unit in a community rehabilitation hospital. New England Journal of Medicine 1990 May 31; 322(22):1572–8.
7. Cohen L. Ottawa's geriatric assessment program "Better care is less expensive". Canadian Medical Association Journal 1989 March 15; 140:710–3.
8. Fretwell MD, Raymond PM, McGarvey ST, Owens N, Traines M, Silliman RA et al. The senior care study: A controlled trial of a consultative/unit based geriatric assessment program in acute care. The Journal of the American Geriatric Society 1990; 38:1073–81.
9. Rubenstein LZ, Wieland GD, Josephson KR, Rosebrook B, Sayre J, Kane RL. Improved survival for frail elderly inpatients on geriatric evaluation unit (GEU): Who benefits? Journal of Clinical Epidemiology 1988; 41(5):441–9.
10. Sloane P. Nursing home candidates: Hospital inpatient trial to identify those appropriately assignable to less intensive care. The Journal of the American Geriatric Society 1980; 28:11.
11. Katz S, Akpom A. A measure of primary sociobiological functions. International Journal of Health Services 1976; 6(3):493–507.
12. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M et al. Development and validation of a geriatric depression screening scale: A preliminary report. Journal of Psychiatric Research 1983; 17(1):37–49.

13. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research 1975; 12:189–98.
14. Romanko L, Pereles L. An evaluation of a pharmacy assessment for geriatric patients. Canadian Journal of Hospital Pharmacy 1992; 45(1):15–20.
15. Engelman L. Stepwise logistic regression by maximum likelihood or asymptotic covariance. BMDP Statisical Software 1980; Technical Report #74.
16. Smith A, Slaughter SE, Berg K. Evaluation of a geriatric rehabilitation unit. Gerontology 1987 Fall; 15–17.
17. Narain P, Rubenstein LZ, Weiland GD, Rosebrook B, Strome LS, Pietruszka PA. Predictors of immediate and 6-month outcomes in hospitalized elderly patients. The Journal of the American Geriatric Society 1988; 37:775–83.
18. Clarfield AM. The reversible dementias: Do they reverse? Annual of Internal Medicine 1988; 476–86.
19. McCusker J, Mundt DJ, Stoddard AM, Cole E, Whitbourne SK, Simmons JE. Outcomes of a geriatric rehabilitation program in a long-term care facility. The Journal Aging and Health 1989 Nov; 485–506.
20. Wachtel TM, Derby C, Fulton JP. Predicting the outcome of hospitalization for elderly persons: Home versus nursing home. Southern Medical Journal 1984; 77(10): 1283–5.
21. Myers JK, Wiessman MM, Tischler GL, Holzer III CE, Leaf PJ, Orvaschel H et al. Six-month prevalence of psychiatric disorders in three communities. Archives of General Psychiatry, 1984 Oct; Vol 41.
22. Wachtel TM, Fulton JP, Goldfarb J. Early prediction of discharge disposition after hospitalization. The Gerontologist 1987; 27(1):98–103.
23. Hood JC, Murphy JE. Patient noncompliance can lead to hospital readmissions. Hospitals 1978; 52:79–84.
24. Beers M, Ouslander J. Risk factors in geriatric drug prescribing: A practical guide to avoiding problems. Drugs 1989; 37:105–12.

The Beck Anxiety Inventory: Psychometric Properties in a Community Population[1]

Augustine Osman,[2,3] Francisco X. Barrios,[2] Diane Aukes,[2] Joylene R. Osman,[2] and Kathleen Markway[2]

Accepted: August 18, 1993

This study presents data on the norms and psychometric properties of the Beck Anxiety Inventory (BAI), using a sample of 225 community adult volunteers. Maximum-likelihood confirmatory factor analyses of previously published models of the BAI: a two-factor model and a five-factor model, showed that the fit of each model was unacceptable. Also, the fit of the single-factor model was poor. Exploratory principal-components analyses with varimax and oblique rotations suggested four BAI components within this sample. Satisfactory levels of reliability were established for the BAI subscales. Finally, the relations between the BAI total and subscale scores and a related measure of anxiety and with another self-report measure of psychological distress were examined.

KEY WORDS: Beck Anxiety Inventory; norms; factor analyses.

INTRODUCTION

In 1988, Beck, Epstein, Brown, and Steer described the development and initial psychometric properties of the Beck Anxiety Inventory (BAI). This 21-item self-report inventory was designed to assess the severity of anxiety symptoms in adults and adolescents. In their initial psychiatric outpatient sample (n = 160), Beck *et al.* (1988) identified two factors for the BAI: somatic and subjective anxiety/panic. The factors showed good internal consistency, test–retest reliability, and convergent/divergent validity.

[1]This paper was presented at the 27th Annual Convention of the Association for Advancement of Behavior Therapy, Atlanta, Georgia, November 1993.

[2]Department of Psychology, Baker 439, University of Northern Iowa, Cedar Falls, Iowa 50614-0505.
[3]To whom correspondence should be addressed.

0882-2689/93/1200-0287$07.00/0

Recently, Borden, Peterson, and Jackson (1991) administered the BAI to 293 introductory psychology college students and identified five factors: subjective fear and anxiety, somatic nervousness, neurophysiological, muscular/motoric, and respiration. Unfortunately, the coefficients alpha were not reported for the factor scales. But, in a second sample of 40 undergraduates, Borden *et al.* (1991) reported moderate and significant correlations between the BAI and several physiological measures and with subjective distress. Results provided support for the concurrent validity of the BAI.

To date, there has been no published study of the norms and psychometric properties of the BAI in a community nonpatient adult sample, despite the noted relevance of nonclinical normative data to clinical assessors (see Kendall & Grove, 1988).

The present study presents data concerning the norms and various psychometric properties of the BAI in an adult community sample. Specifically, the factor structure and internal consistency reliability of the BAI were evaluated. Finally, we examined the relations between the BAI and a related measure of anxiety as well as with a self-report measure of psychological distress.

METHOD

Subjects

Subjects were 225 adult volunteers: 66 men (age: M = 36.2 years, SD = 11.9 years; range, 20–72 years) and 159 women (age: M = 37.1 years, SD = 12.0 years; range, 20–74 years) recruited from a moderate-size Midwestern community. The sample was predominantly White and included individuals with a range of background education: partial junior high to graduate studies. Sixty-four percent were married and 36% were unmarried (including those never married, divorced, separated, and widowed). Each subject provided informed consent to participate.

Measures and Procedure

Each participant completed a packet of self-report measures that included a background information questionnaire, the BAI, and two other measures with established reliability and validity.

The Beck Anxiety Inventory (BAI; Beck et al., *1988).* The BAI consists of 21 items that assess the severity of anxiety. Subjects rate each item, using a 4-point scale: 0 (*not at all*) to 3 (*severely, I could barely stand it*). Anxiety severity is the total raw-score sum across all 21 items. The total scores range from 0 to 63.

The Cognitive-Somatic Anxiety Questionnaire (CSAQ; Schwartz, Davidson, & Goleman, 1978). The CSAQ is a 14-item self-report measure of the cognitive and somatic symptoms of anxiety. Each item is rated on a 5-point scale ranging from 1 (*not at all*) to 5 (*very much*). Satisfactory reliability and validity have been established for the CSAQ. This measure was included in 158 (68.7%) of the questionnaire packets.

The Brief Symptom Inventory (BSI; Derogatis, 1992). The BAI is a short form of the Symptom Checklist 90R (SCL-90R; Derogatis, 1983). This 53-item inventory is a measure of current symptoms of psychological distress. Each item is rated on a 5-point scale in Likert format ranging from 0 (*not at all*) to 4 (*extremely*). Separate subscales can be scored that measure nine dimensions, as well as three global indices of psychological distress.

Residential areas within each of the four zones of the Cedar Falls community were randomly selected from the city zone map provided by the Department of Planning and Zoning for questionnaire distribution. Undergraduate senior research assistants distributed (door-to-door) the questionnaire packets. The completed questionnaires were collected within 48 hr of administration. Of the 230 homes contacted, only 5 returned incompleted questionnaires to the research assistants. No data regarding income and occupation were obtained.

RESULTS AND DISCUSSION

Factor Analyses

Maximum-likelihood confirmatory factor analyses (LISREL 7; Joreskog & Sorbom, 1990) were used to assess the generalizability of the two-factor model (Beck *et al.*, 1988) and the five-factor model (Borden *et al.*, 1991) to our community adult data. Additionally, the adequacy of the BAI as a unidimensional self-report measure of anxiety was assessed by forcing all items to load on a single factor.

Because of problems related to the use of the chi-square statistic, such as sensitivity to large sample size (see Cole, 1987), we used three practical measures of fit: the goodness-of-fit (GFI, values greater than .90), the adjusted goodness-of-fit (AGFI, values greater than .80), and the root mean square residual (RMS, values less than .10) to evaluate the adequacy of each model.

The GFI = .72, the AGFI = .65, and the RMS = .12 values for the two-factor model were not an adequate fit to the observed data. Similarly, for the one-factor model, the GFI = .56, and the AGFI = .48 values

suggested that this model did not fit the data adequately. Finally, all indices of fit also suggested that the five-factor model (GFI = .81; AGFI = .75; RMS = .08) approached but did not provide a satisfactory fit to the data for our sample.

Because of the poor solutions obtained, exploratory principal-components analyses with varimax and oblique rotations were conducted to explore the factor structure of the BAI in this sample. A high Kaiser–Meyer–Olkin ratio (KMO = .90) suggested that the principal-components analysis was appropriate. Both the eigenone criterion and the scree plot suggested a four-factor model. Items loading .45 or above on a primary factor and not on another factor were retained. Varimax and oblique rotations resulted in similar results. Results of the varimax procedure are shown in Table I. The four factors accounted for 64.6% of the total variance.

Reliability and Subscale Correlations

Coefficients alpha for the BAI total and subscale scores were high and acceptable (see Table II). All item-subscale correlations (values were .70 or greater) were also adequate. The pattern of correlations among the subscales was similar for men and women. For men, however, the alpha value for the panic factor was low.

Norms and Subscale Validity

Normative data for the study sample are presented in Table II. Initial analysis (see Table II) showed negative and significant relationships between age and scores on the BAI subjective ($r = -.25$, $p < .01$), autonomic ($r = -.17$, $p < .05$), and panic ($r = -.14$, $p < .05$) subscales, suggesting that younger subjects may score higher than older subjects on these subscales. A 2 (Gender) × 4 (Subscale) multivariate analysis of covariance (covariate, age) showed significant differences between the genders [$F(4,219) = 3.17$, Hotelling's $T^2 = .06$, $p < .01$]. As shown in Table II, women scored significantly higher than men on the panic ($t = 2.77$, $p < .01$) and autonomic ($t = 2.84$, $p < .01$) subscales as well as on the BAI total score ($t = 2.58$, $p < .05$).

Finally, descriptive discriminant analysis, undertaken to examine gender differences across the 21 items, showed that 9 of the 21 items (see Table I) were useful in accounting for subgroup separation.

Table I. Sorted Rotated Factor Loadings and Communalities (h^2) from Principal Components

Item		Factor loading 1	2	3	4	h^2
Factor 1.	Subjective					
10.	Nervous	.83	.09	.21	.01	.74
17.	Scared[a]	.79	.12	.18	.15	.70
5.	Worst happening[a]	.78	.18	.14	.07	.67
14.	Losing control	.77	.21	.15	.15	.68
4.	Unable to relax	.75	.13	.22	−.03	.63
9.	Terrified[a]	.75	.23	.12	.20	.66
Factor 2.	Neurophysiological					
13.	Shaky	.19	.78	.27	.11	.73
8.	Unsteady[a]	.28	.76	.24	.11	.72
3.	Wobbliness	.20	.75	.14	.02	.63
19.	Faint[a]	−.01	.73	.20	.18	.61
6.	Dizzy	.09	.71	.15	.23	.59
1.	Numbness	.14	.70	−.05	−.06	.51
12.	Trembling	.25	.69	.40	.09	.71
Factor 3.	Autonomic					
20.	Face flushed[a]	.21	.23	.78	.08	.72
2.	Feeling hot[a]	.12	.23	.77	.09	.67
21.	Sweating	.27	.28	.68	.13	.63
18.	Indigestion	.34	.08	.52	.01	.39
Factor 4.	Panic					
11.	Choking	−.05	−.02	.19	.80	.68
15.	Breathing	.17	.32	.23	.75	.75
16.	Dying[a]	.32	.14	−.23	.70	.66
7.	Heart pounding	.37	.25	.35	.38	.47
Eigenvalue		8.12	2.42	1.60	1.41	
Variance (%)		38.7	11.5	7.6	6.7	

[a] Item discriminated between male and female subjects.

Table II. Means, Standard Deviations, Alpha Values, and Subscale Intercorrelations

Variables	BAI subscales 1	2	3	4	BAI total	M	SD	α
Men (n = 66)								
1. Subjective	1.00					4.27	3.76	.86
2. Neurophysiological	.21	1.00				1.83	2.69	.86
3. Autonomic	.27[a]	.38[b]	1.00			1.77	2.01	.75
4. Panic	.34[b]	.32[b]	.33[b]	1.00		.96	1.41	.58
5. BAI total	.77[b]	.68[b]	.65[b]	.61[b]	1.00	8.83	6.92	.85
Women (n = 159)								
1. Subjective	1.00					5.18	4.81	.91
2. Neurophysiological	.49[b]	1.00				2.86	4.19	.90
3. Autonomic	.57[b]	.55[b]	1.00			2.85	2.79	.78
4. Panic	.48[b]	.47[b]	.39[b]	1.00		1.77	2.22	.71
5. BAI total	.85[b]	.82[b]	.78[b]	.68[b]	1.00	12.66	11.20	.92
Total sample (N = 225)								
1. Subjective	1.00					4.92	4.54	.90
2. Neurophysiological	.45[b]	1.00				2.56	3.84	.89
3. Autonomic	.52[b]	.53[b]	1.00			2.53	2.63	.78
4. Panic	.46[b]	.46[b]	.40[b]	1.00		1.53	2.05	.70
5. BAI total	.83[b]	.80[b]	.76[b]	.68[b]	1.00	11.54	10.36	.92
6. Age	−.25[b]	−.04	−.17[a]	−.14[a]	−.20[b]	36.91	11.98	—
7. Gender	.09	.12	.18[a]	.18[a]	.17[a]	—	—	—

[a] $p < .05$. [b] $p < .01$.

Relation of the BAI to the CSAQ

Zero-order and partial correlation coefficients were computed between the BAI and the CSAQ subscales. All zero-order correlations between the BAI subscales and the CSAQ-Cognitive (see Table III) were positive and significant for men, women, and the total sample. Similarly, for women and the total sample, all BAI subscales correlated positively and significantly with the CSAQ-Somatic. But for men, only the BAI neurophysiological and autonomic subscales correlated with the CSAQ-Somatic.

When we statistically controlled for general psychological distress (Global Severity Index; GSI), the magnitudes of these correlations dropped slightly to markedly for both genders. Also, for both genders and total sample, partialling out GSI had the greatest impact on the relationships between the BAI subscales and the CSAQ-Cognitive. Only the BAI subjective remained correlated significantly with the CSAQ-Cognitive for women and the total sample. Further examination of the relationships between the CSAQ-Somatic and the BAI subscales showed that, for men, the CSAQ-Somatic maintained its significant relationships with the neurophysiological and autonomic subscales when GSI was covaried. For women, the CSAQ-Somatic remained correlated with only two of the four BAI subscales. Results provide initial support for the concurrent validity of the BAI in this sample.

Relation of the BAI to Psychological Distress

Pearson product-moment correlation coefficients were computed for the total sample and for men and women separately. Results are presented in Table IV. For women, all four BAI subscales significantly correlated with the GSI and all nine BSI subscales. These findings do not support the specificity of the BAI with the BSI in this subsample. For the men, the BAI subjective subscale correlated positively and significantly with the GSI and all nine BSI subscales. Further examination of the pattern of correlations showed that the BAI neurophysiological, autonomic, and panic subscales correlated as highly with the BSI anxiety subscale as with the remaining BSI subscales. Similarly, the pattern of correlations between the BAI and the BSI does not provide strong evidence for the specificity of the BAI for men.

Table III. Correlations (Zero-Order and Partial) of the BAI with the CSAQ

Variable	BAI subscales				BAI total score
	Subjective	Neurophysiological	Autonomic	Panic	
Men (n = 54)					
CSAQ-Cognitive	.44[c]	.33[c]	.30[c]	.33[c]	.51[c]
Partial r[a]	.15	.24	.16	.04	.26[b]
CSAQ-Somatic	.17	.31[b]	.44[c]	.24	.38[c]
Partial r[a]	.00	.26[b]	.39[b]	.12	.29[b]
Women (n = 104)					
CSAQ-Cognitive	.57[c]	.27[c]	.32[c]	.39[c]	.52[c]
Partial r[a]	.32[c]	−.05	.03	.08	.18
CSAQ-Somatic	.38[c]	.49[c]	.24[c]	.44[c]	.50[c]
Partial[a]	.12	.33[c]	−.01	.24[c]	.27[c]
Total subsample (n = 158)					
CSAQ-Cognitive	.52[c]	.29[c]	.32[c]	.36[c]	.51[c]
Partial[a]	.25[c]	.07	.11	.10[c]	.23[c]
CSAQ-Somatic	.31[c]	.43[c]	.31[c]	.38[c]	.46[c]
Partial[a]	.11	.32[c]	.17[b]	.23[c]	.30[c]

[a] Controlling for general psychological distress (GSI).
[b] $p < .05$ (two-tailed test).
[c] $p < .01$ (two-tailed test).

Table IV. Correlations of the BAI with the Brief Symptom Inventory (BSI)[a]

	BAI Subscale									
	SOM	OC	INT	DEP	ANX	HOS	PHOB	PAR	PSY	GSI
Men (n = 66)										
Subjective	.39[c]	.39[c]	.50[c]	.63[c]	.69[c]	.50[c]	.42[c]	.49[c]	.60[c]	.68[c]
Neurophysiological	.44[c]	.27[b]	.25[b]	.22	.16	.23	.02	.15	.16	.26[b]
Autonomic	.44[c]	.43[c]	.19	.27[b]	.23	.33[c]	.12	.29[c]	.19	.35[c]
Panic	.37[c]	.42[c]	.33[c]	.28[b]	.37[c]	.13	.25[b]	.34[c]	.24[b]	.39[c]
BAI Total	.59[c]	.53[c]	.49[c]	.57[c]	.58[c]	.48[c]	.32[c]	.48[c]	.49[c]	.66[c]
Women (n = 159)										
Subjective	.55[c]	.61[c]	.57[c]	.70[c]	.72[c]	.58[c]	.56[c]	.61[c]	.71[c]	.71[c]
Neurophysiological	.66[c]	.44[c]	.40[c]	.50[c]	.48[c]	.35[c]	.47[c]	.45[c]	.46[c]	.51[c]
Autonomic	.65[c]	.51[c]	.50[c]	.54[c]	.55[c]	.44[c]	.45[c]	.49[c]	.52[c]	.58[c]
Panic	.48[c]	.39[c]	.39[c]	.49[c]	.46[c]	.36[c]	.45[c]	.45[c]	50[c]	.49[c]
BAI Total	.74[c]	.63[c]	.60[c]	.72[c]	.72[c]	.56[c]	.62[c]	.64[c]	.71[c]	.74[c]
Total sample (n = 225)										
Subjective	.50[c]	.55[c]	.54[c]	.67[c]	.68[c]	.56[c]	.50[c]	.57[c]	.68[c]	.69[c]
Neurophysiological	.59[c]	.39[c]	.35[c]	.42[c]	.38[c]	.32[c]	.34[c]	.38[c]	.40[c]	.44[c]
Autonomic	.58[c]	.47[c]	.40[c]	.46[c]	.43[c]	.41[c]	.33[c]	43[c]	.44[c]	49[c]
Panic	.44[c]	.37[c]	.35[c]	.42[c]	.39[c]	.30[c]	.36[c]	.42[c]	.44[c]	.43[c]
BAI Total	.68[c]	.58[c]	.54[c]	.66[c]	.63[c]	.53[c]	.51[c]	.59[c]	.65[c]	.68[c]

[a] SOM, somatization; OC, obsessive–compulsive; INT, interpersonal sensitivity; DEP, depression; ANX, anxiety; HOS, hostility; PHOB, phobic anxiety; PAR, paranoid ideation; PSY, Psychoticism; GSI, general severity index.
[b] $p < .05$ (two-tailed test).
[c] $p < .01$ (two-tailed test).

In summary, the present study presents preliminary data on the norms and psychometric properties of the BAI for a sample of community-dwelling adults. Our normative data may be useful for interpreting meaningfully the responses of clinically anxious subjects (see Kendall & Grove, 1988).

Results of the confirmatory and exploratory factor analyses suggest that the factor structure of the BAI in our sample may be different from those reported for a mixed diagnosed sample of psychiatric outpatients (Beck *et al.*, 1988) and college students (Borden *et al.*, 1991). The extracted factors showed acceptable levels of reliability. Age correlated significantly with the BAI subjective, autonomic, and panic subscales, suggesting consideration of this factor in the use of the BAI. Also, two of the BAI subscales were useful in differentiating between male and female subjects in this sample. For the total sample, the mean score (M = 11.54) on the total BAI was slightly higher than those reported by Borden *et al.* (1991) for the college sample (M = 10.75). The mean score for Beck and co-workers' (1988) outpatient psychiatric sample (M = 22.35) was higher than the mean score obtained in our sample. Consistent with the findings of previous studies, female subjects scored significantly higher than male subjects on the total BAI.

Correlations between the BAI subscales and a related measure of anxiety provided initial evidence for the concurrent validity of the BAI. However, correlations between the BAI subscales and a self-report measure of psychological distress suggested that the symptoms of anxiety reported by men or women were difficult to differentiate from their self-reports of general psychological distress. The lack of specificity of the BAI with the BSI may relate in part from the moderate to high correlations found (range, r = .43 to .77) among the nine BSI subscales in this study. Also, as one reviewer pointed out, there is additional evidence that suggests that the SCL-90R subscales lack discriminant validity (e.g., see Cyr, Doxey, & Vigna, 1988). Because the BSI items were drawn directly from the SCL-90R (see Derogatis, 1992), the BSI subscales may lack discriminant validity. Future researchers might examine the relationships between the BAI and validated measures that assess specific psychopathology such as the State-Trait Anxiety Inventory (STAI, Form Y; Spielberger, 1983) and the Cognition Checklist (CCL; Beck, Brown, Steer, Eidelson, & Riskind, 1987) to assess the specificity of the BAI. In general, our findings suggest that the BAI is a promising multidimensional measure of anxiety in this sample.

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers for their constructive comments on an early version of the manuscript.

REFERENCES

Beck, A. T., Brown, G., Steer, R. A., Eidelson, J. I., & Riskind, J. H. (1987). Differentiating anxiety and depression: A test of the cognitive-specificity hypothesis. *Journal of Abnormal Psychology, 96,* 179-183.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893-897.

Borden, J. W., Peterson, D. R., & Jackson, E. A. (1991). The Beck Anxiety Inventory in nonclinical samples: Initial psychometric properties. *Journal of Psychopathology and Behavioral Assessment, 13,* 345-356.

Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology, 55,* 584-594.

Cyr, J. J., Doxey, N. C. S., & Vigna, C. M. (1988). Factorial composition of the SCL-90—R. *Journal of Social Behavior and Personality, 3,* 245-252.

Derogatis, L. R. (1983). *SCL-90R: Administration, scoring and procedures manual II.* Townson, MD: Clinical Psychometrics Research.

Derogatis, L. R. (1992). *The Brief Symptom Inventory (BSI): Administration, scoring and procedures manual II.* Townson, MD: Clinical Psychometrics Research.

Joreskog, K. G., & Sorbom, D. (1990). LISREL 7: *Estimation of Linear Structural Equation Systems.* Mooresville, IN: Scientific Software.

Kendall, P. C., & Grove, W. M. (1988). Normative comparisons in therapy outcome. *Behavioral Assessment, 10,* 147-158.

Schwartz, G. E., Davidson, R. J., & Goleman, D. J. (1978). Patterning of cognitive and somatic processes in the self-regulation of anxiety: Effects of meditation versus exercise. *Psychosomatic Medicine, 40,* 321-328.

Spielberger, C. D. (1983). *Manual for the State-Trait Inventory (Form Y).* Palo Alto, CA: Consulting Psychologists Press.