Behavior Research Methods, Instruments, & Computers
1985, 17 (4), 445-449

METHODS & DESIGNS

Reliability and validity of portable blood pressure/pulse units during a competitive challenge

MICHAEL J. FOLLICK, DAVID K. AHERN, and LARRY GORKIN
Brown University, Providence, Rhode Island
and The Miriam Hospital, Providence, Rhode Island

The aim of the present study was to determine the reliability and validity of recordings from three portable blood pressure/pulse units compared with simultaneous recordings from a polygraph used as the standard for these comparisons. The units tested included the Sears Digital Z153, Industrial Biomedical Sensor (IBS) Digital SD-700, and the Sphygmostat B-350/PU-102 (analog scale). Thirty healthy subjects were exposed to Pac-Man video game and mental arithmetic stressors. Each stressor was preceded by relaxation and followed by recovery periods. During each period, systolic/diastolic blood pressure and pulse were taken from one of the three portable units and the polygraph. The three units were reliable in recording systolic and diastolic blood pressure and heart rate across baseline, stressor, and recovery periods. The blood pressure/pulse readings from the digital units by Sears and IBS corresponded more closely to simultaneous polygraph recordings than did the analog scale device by Sphygmostat. These results suggest that the digital blood pressure/pulse units employed in this investigation are reliable and valid measurement devices suitable for field research with clinical populations.

Numerous investigations have demonstrated that laboratory stressors can produce elevations in blood pressure (Falkner, Kushner, Onesti, & Angelakos, 1981; Obrist, 1981). There is growing concern, however, that laboratory stressors are not representative of naturally occurring stressors which may contribute to the onset or progression of heart disease (Light & Obrist, 1983). Accordingly, there is interest in conducting field research to further evaluate these relationships and improve understanding of the role of stress in hypertension. Recent research with ambulatory blood pressure monitoring devices indicates that home-based readings are more reliable measures of pressure elevations than office readings (Kleinert et al., 1984), and blood pressure changes during stressful situations in the natural environment (e.g., work) are better predictors of complications from hypertension than measurements taken in a physician's office (Devereux et al., 1983).

Invasive measurements (e.g., from arteriocatheterization) or polygraph recordings of blood pressure and heart rate under stress conditions are difficult if not impossible to obtain in the natural environment. While noninvasive ambulatory blood pressure recording devices may be suitable alternatives, they are expensive and are often cumbersome for the patient. Portable automatic blood pressure/pulse units may provide a preferable alternative when continuous monitoring is not essential, but the adequacy of these devices has yet to be established.

Prior research evaluating the performance of selected automatic devices for measuring blood pressure has shown that earlier generation instruments are inadequate for use in clinical trials (Labarthe, Hawkins, & Remington, 1973). However, the recent emphasis on high blood pressure detection and monitoring has resulted in a new generation of portable automatic blood pressure monitors which are affordable, easy to use, and constructed with more advanced solid state circuitry and transducers. The use of these newer portable blood pressure/pulse units in field research is warranted only if they demonstrate acceptable levels of reliability and validity. Consequently, the present study was designed to determine the reliability and validity of three commonly employed blood pressure/pulse units. One of the devices selected for evaluation is currently being used in a multicenter national clinical trial, the Cardiac Arrhythmia Pilot Study (CAPS), to examine the relationship between cardiovascular reactivity to stress and the frequency and grade of arrhythmias in post-myocardial infarction patients. Two additional units were selected for comparison because of their price range, available features, and the fact that they have been used in clinical research. Polygraph recordings of blood pressure and heart rate were taken for purposes of comparison with the readings of automatic portable units.

This research was supported in part by contract award No. N01-HU-28010, Cardiac Arrhythmia Pilot Study, Clinical Trials Branch, NHLBI.
Please address reprint requests to: M. J. Follick, The Miriam Hospital, Division of Behavioral Medicine, Providence, RI 02906.

METHOD

Subjects

Thirty healthy subjects with no documented cardiovascular disease were recruited. The sample consisted of 18 males and 12 females who ranged in age from 14 to 65 years. Subjects were paid $5 for their participation in the study.

Instrumentation

Three portable blood pressure/pulse units were used: Sears Digital Z153 ($200), Industrial Biomedical Sensor (IBS) Digital SD-700 ($700), and the Sphygmostat B-350/PU-102 ($300, analog scale). Each of the portable units contains a microphone attached to an occluding cuff that is inflated manually. The units were not altered prior to testing; they were calibrated, presumably, at the factory. The Sears and Sphygmostat units detect systole by the onset of blood flow through the artery and detect diastole by the shift in blood flow (i.e., muffling). The IBS unit derives systole and diastole by the oscillometric method based on an average blood pressure. The Sears and IBS units derive pulse rate based on the number of Korotkoff beats within a given time frame converted to beats per minute. The Sphygmostat detects pulse rate based on a finger phototransducer, which converts the change in the quantity of light to an electrical signal in beats per minute.

Systolic and diastolic blood pressures (SBP/DBP) were also monitored with a Grass Model 7 polygraph with a 7P8 preamplifier and a 1010 microphone (without cuff) to detect Korotkoff sounds. Heart rate (HR) was derived from 5-lead electrocardiographic (ECG) recording on a Grass Model 7 polygraph. Five Con-Med silver leads (LP72-005) attached to silver/silver chloride pregelled electrodes (105-5005) were connected to a Grass 7P4 preamplifier to record ECG. When calibrated for blood pressure, the polygraph was found to be strongly related to a mercury column through the range of 50-mm Hg to 240-mm Hg [$r(11) = .999$, $p < .001$], with no absolute value difference greater than 1.5% between the two measurement devices.

Procedure

Subjects were assigned randomly to one of three portable blood pressure/pulse units (Sears, IBS, or Sphygmostat) with the restriction that 10 subjects were assigned to each unit. Simultaneous blood pressure and ECG recordings were monitored by the polygraph. Thus, each subject had assessments conducted by only one portable unit and the polygraph. Blood pressures were assessed by separate microphones for the portable unit and the polygraph. Both microphones were placed adjacent to the brachial artery of the nondominant arm, with the microphone connected to the portable unit located medial to the microphone connected to the polygraph. The microphones were placed approximately 1 cm apart, allowing the occluding cuff of the portable unit to cover the taped-on microphone of the polygraph.

Physiological monitoring occurred during two experimental tasks: a 4-min Pac-Man videogame and a 4-min mental arithmetic stressor. Pac-Man was played on the Atari 5200 system with a 9-in. color television screen. The Pac-Man joystick was secured to a hospital tray with Velcro®. This procedure allowed for operation of the joystick with the dominant hand; the nondominant hand (from which readings were obtained) was restrained by Velcro. The mental arithmetic problem consisted of subtracting serial 7s from 2193 while being challenged periodically by the experimenter to respond more quickly (e.g., "please go faster," " try to go as fast as you can"). Each task was preceded by a 6-min tape of passive relaxation instructions to stabilize the physiological measures and was followed by a 4-min recovery period. The Pac-Man game always preceded mental arithmetic.

Cardiovascular measures were monitored twice per experimental phase, except ECG was continuous. More specifically, assessments were taken prior to and immediately following the relaxation tape, and at 2 min and 4 min following initiation of either task or recovery period. Accordingly, there were 12 samples of SBP and DBP taken for each subject.

Data Recording

Blood pressure and pulse measures were provided by LED readouts for both the Sears and IBS units. Readings from the Sphygmostat unit were more difficult to obtain and involved the following operations: (1) the unit signaled systole with a flashing light and tone; (2) the analog scale was read for SBP; (3) the analog scale was read for pulse; (4) diastole was signaled by the last flash of light and tone; and (5) the analog scale was read for DBP.

The 5-lead ECG polygraph recordings were converted to instantaneous heart rate (HR) in beats per minute (bpm) and averaged for five consecutive beats around the point of systole to derive the HR level during each experimental phase. The selection of five consecutive beats was arbitrary, although the average of five beats was deemed an adequate representation of rate. Phase IV (muffling) was selected as the criterion for diastolic pressure, as recommended by Steptoe (1980). Muffling produces a qualitative shift in the amplitude of Korotkoff spikes superimposed on the cuff-deflation curve.

The tachographic recordings and polygraph blood pressure recordings were scored by hand for heart rate, systole, and diastole. All of the protocols were scored by one rater, who was not informed of the assignment of portable units to subjects, but was aware of the sequence of experimental phases. Independently, a second blind rater scored 10% of the protocols chosen at random as a reliability check. Results indicated strong coefficients of interrater reliability for the scoring of polygraph recordings [SBP, $r(36) = .93$, $p < .001$; DBP, $r(36) = .97$, $p < .001$; pulse, $r(36) = .98$, $p < .001$].

Data Analysis

Portable unit reliability was determined by the correlations of adjacent SBP, DBP, and pulse readings per

phase (i.e., the two readings within each phase) and per unit during the Pac-Man task.[1] Although reliability coefficients of psychological measures are often expected to be .90 or higher, this criterion does not appear warranted with psychophysiological measures. Arena, Blanchard, Andrasik, Cotch, and Myers (1983) reported reliability coefficients of .50 to .60 for absolute values of HR assessed twice during mental arithmetic on 15 subjects, 1 week apart. With a larger subject sample (n = 67), Sher, Walitzer, Mannion, and Hammer (1984) reported similar results for interbeat intervals recorded during an interpersonal stressor assessed 1 week apart. Williamson, Waters, Bernard, Faulstich, and Blouin (1985) reported HR correlations of .52 for baseline and .20 for an oral "intelligence quiz" assessed over a 2-week interval. SBP and DBP coefficients ranged from .25 to .56 for this component of the Williamson et al. study. While these results point to only modest cardiovascular reliability coefficients gathered 1 to 14 days apart, a more conservative estimate (r = .70) was used in the present study due to the analyses of assessments taken minutes apart.

In measuring validity, direct comparisons of the portable units and the polygraph were not possible because the three units were not tested simultaneously (i.e., subjects were assigned to only one unit and the polygraph). Instead, validation of the portable units was tested by two analyses: (1) correlations of the portable unit's second readings with analogous polygraph recordings of SBP, DBP, and HR across varying experimental conditions involving both tasks using a convergent validity criterion of .70 (cf. Campbell & Fiske, 1959) and (2) unit × phase repeated measures ANOVA of absolute mean SBP, DBP, and pulse/HR difference scores between the polygraph and each portable unit across experimental conditions during mental arithmetic.[2] Absolute scores were selected to preclude discrepancies in opposite directions, between each unit and analogous polygraph recordings, which would act to cancel one another and thereby artificially inflate the resultant validation coefficients. For the analyses, mean values were substituted for missing data due to movement artifact. Substitutions occurred for two pulse readings from the Sears unit, one DBP reading from the IBS unit, and two blood pressure and one pulse readings from the Sphygmostat unit. To determine whether the tasks, as expected, produced higher levels of cardiovascular reactivity than their respective baselines, and whether the tasks differed from one another in terms of cardiovascular reactivity, two-way (unit × phase) repeated measures ANOVAs were conducted separately on SBP, DBP, and HR generated by the portable units.

RESULTS

Correlation coefficients for adjacent readings revealed that 24 of 27 values (8 of 9 values for each unit) were above a conservative lower limit for test-retest reliability (i.e., r = .70). These results are presented in Table 1. The coefficients obtained below this cutoff were for the Sears unit's baseline pulse [r(10) = .63], IBS unit's Pac-Man task DBP [r(9) = .56], and the Sphygmostat unit's Pac-Man task pulse [r(10) = .67]. (One subject was eliminated from the IBS Pac-Man DBP analysis due to evidence of an outlier score; his first blood pressure reading for this parameter was 170/078, an unlikely difference of 92-mm Hg between SBP and DBP, and placement on the graph was distant from the other nine subjects assessed twice for DBP during the Pac-Man stressor.) Thus, the pattern of results appears to support the claim that reliability was acceptable for each unit.

Table 1
Correlations of Adjacent SBP, DBP, and Pulse per Phase by Unit During the Pac-Man Procedure

Unit (n = 10)*	BL	P-M	RC
	SBP (mm Hg)		
Sears	.92†	.94†	.98†
IBS	.78†	.86†	.87†
Sphygmo	.94†	.87†	.81†
	DBP (mm Hg)		
Sears	.95†	.79†	.85†
IBS	.93†	.56	.82†
Sphygmo	.97†	.87†	.89†
	Pulse (bpm)		
Sears	.63‡	.86†	.83†
IBS	.86†	.79†	.94†
Sphygmo	.90†	.67‡	.94†

*Note—Sphygmo = Sphygmostat; SBP = systolic blood pressure; DBP = diastolic blood pressure; Pulse = pulse rate; BL = baseline; P-M = Pac-Man; RC = recovery; mm Hg = millimeters of mercury; bpm = beats per minute. *Analyses were based on 10 subjects per unit, except for the IBS during Pac-Man when only 9 subjects were used. †$p < .01$. ‡$p < .05$.*

The initial validation test consisted of correlating portable unit readings with analogous polygraph recordings. Results are presented in Table 2. The Sphygmostat unit had 8 of 18 coefficients below the .70 validation criterion,

Table 2
Correlations of Portable Units to Polygraph for SBP, DBP, and HR Across Varying Experimental Conditions (Second Recordings)

Unit (n = 10)	BL I	P-M	RC I	BL II	MA	RC II
			SBP (mm Hg)			
Sears	.78*	.93*	.83*	.97*	.96*	.95*
IBS	.47	.91*	.90*	.90*	.94*	.79*
Sphygmo	.74*	.83*	.67†	.79*	.95*	.42
			DBP (mm Hg)			
Sears	.92*	.97*	.95*	.98*	.95*	.99*
IBS	.94*	.96*	.95*	.85*	.75*	.72†
Sphygmo	.81*	.77*	.75*	.74*	.92*	.38
			HR (bpm)			
Sears	.96*	.96*	.96*	.90*	.94*	.97*
IBS	.96*	.83*	.89*	.92*	.92*	.91*
Sphygmo	.19	.43	.17	.14	.39	.03

*Note—Sphygmo = Sphygmostat; SBP = systolic blood pressure; DBP = diastolic blood pressure; HR = heart rate; BL I = first baseline; P-M = Pac-Man; RC I = first recovery; BL II = second baseline; MA = mental arithmetic; RC II = second recovery; mm Hg = millimeters of mercury; bpm = beats per minute. *$p < .01$. †$p < .05$.*

Table 3
Unit by Phase Absolute Mean Differences and Standard Deviations (SD) Between Polygraph and Portable Units During Mental Arithmetic

	BL		MA		RC	
Unit (n = 10)	Mean	SD	Mean	SD	Mean	SD
			SBP (mm Hg)			
Sears	5.90	3.36	4.65	2.15	3.05	1.24*
IBS	3.95	2.01	5.90	3.30	5.45	3.63
Sphygmo	7.10	3.56	6.10	3.46	10.60	9.01*
			DBP (mm Hg)			
Sears	3.55	2.02	5.40	5.34	3.10	2.64
IBS	3.40	2.47	4.20	2.65	3.65	3.22
Sphygmo	4.50	5.09	5.35	5.83	7.10	8.21
			HR/Pulse (bpm)			
Sears	3.40	1.45	3.80	2.08	3.50	1.89
IBS	2.50	1.18	6.50	3.84	4.25	4.28
Sphygmo	6.80	15.10	9.25	14.07	7.45	15.95

*Note—Sphygmo = Sphygmostat; SBP = systolic blood pressure; DBP = diastolic blood pressure; HR = heart rate; Pulse = pulse rate; BL = baseline; MA = mental arithmetic; RC = recovery. *p < .05.*

whereas there were none for the Sears unit and only one anomalous coefficient for the IBS unit. For the Sphygmostat unit, the coefficients below criterion included all six pulse/HR comparisons as well as mental arithmetic recovery for SBP and DBP. The one IBS unit coefficient below criterion involved baseline SBP for Pac-Man.

The second test of validation involved a series of repeated measures ANOVAs of absolute difference scores for SBP, DBP, and HR unit readings and simultaneous polygraph recordings during Pac-Man and mental arithmetic. Results for mental arithmetic only are presented in Table 3. Of six analyses (2 types of task × 3 cardiovascular parameters) involving 3 units × 3 experimental levels (baseline, task, recovery), only the analysis of the absolute mental arithmetic SBP was significant. For these analyses, there were effects for unit [$F(2,27) = 3.49$, $p < .05$], and unit by phase [$F(4,54) = 3.69$, $p < .01$]. However, Tukey posttests revealed no differences between groups for the main effect and revealed only one significant difference in a follow-up analysis of unconfounded means for the interaction (cf. Cicchetti, 1972). The Sears unit (mean = 3.05) was more concordant with the polygraph than the Sphygmostat unit was (mean = 10.60) during recovery from mental arithmetic. All 3 units tended to be within 3 to 10 absolute mean units of the respective polygraph readings, although the Sphygmostat tended to be higher, albeit nonsignificantly, across most comparisons.

The test of different levels of cardiovascular arousal as a function of experimental phase revealed no significant interaction of unit × phase for either SBP [$F(6,78) = 1.11$, $p > .05$], DBP [$F(6,78) = 1.54$, $p > .05$], or pulse [$F(6,81) < 1$]. Collapsing across units, however, there was a main effect for phase for SBP [$F(3,78) = 43.59$, $p < .001$], DBP [$F(3,78) = 26.57$, $p < .001$], and pulse [$F(3,81) = 14.82$, $p < .001$]. These results are presented in Table 4. Tukey posttests indicated that both Pac-Man and mental arithmetic generated significantly greater arousal than their respective baselines in all three cardiovascular parameters, except Pac-Man pulse rate versus baseline pulse rate. There were no differences between the two tasks on any of the measures. The experimental manipulations resulted in significant and expected changes in the physiological parameters for all units, with no interaction by unit type.

DISCUSSION

The specific instruments examined, the Sears and IBS digital models and the Sphygmostat analog model, showed similar patterns of reliability coefficients for systole, diastole, and pulse across baseline, stress, and recovery phases. Validity coefficients, representing the relationship between measures from the portable units and polygraph recordings, showed greater variability. The Sears and IBS units had consistently high positive correlations with the polygraph except for the IBS SBP during the initial baseline. The latter finding appears surprising, but the IBS is the only unit which works by the oscillometric method. This method makes the IBS unit best suited to detect stronger vibrations or deflections at systole, which are more likely to occur during a stressor than at rest. Hence, one would likely find greater validity for the IBS unit assessing SBP during Pac-Man than at baseline. In contrast, the validity coefficients for the Sphygmostat unit varied widely with very low correlations for heart rate under all conditions and for SBP/DBP during the second recovery phase.

In addition to the differences in validity among the instruments, there were significant discrepancies in performance characteristics. Although the Sears and IBS units were about equal in performance, the IBS unit produced fewer errors than did the Sears unit, thus necessitating

Table 4
Absolute Mean and Standard Deviations (SD) of SBP, DBP, and Pulse Levels by Experimental Phase Across Two Tasks

	BL I		PM		BL II		MA	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
SBP (mm Hg)*	119.72	14.47	128.21	15.12†	116.24	13.27	130.95	16.80†
DBP (mm Hg)*	79.64	12.35	85.40	10.59†	79.38	11.82	89.52	14.80†
Pulse (bpm)‡	77.18	11.10	82.27	10.56	75.60	10.98	85.97	11.99†

*Note—SBP = systolic blood pressure; DBP = diastolic blood pressure; Pulse = pulse rate; BL I = first baseline; PM = Pac-Man; BL II = second baseline; MA = mental arithmetic; mm Hg = millimeters of mercury; bpm = beats per minute. *n = 29. †significantly different from baseline, p < .05. ‡n = 30.*

fewer repeat trials to obtain readings. The Sphygmostat unit, however, did not perform as well as the other two units: (1) the Sphygmostat produced a considerably larger number of undetected samples; (2) the Sphygmostat does not generate error messages, so there is no indication of existing problems during recording; (3) the Sphygmostat is more complex to operate than the other units because of the requirement to monitor a flashing light and tone in combination with an analog scale to obtain readings; and (4) there are different analog scales for blood pressure and pulse rate.

Assuming that the Sears and IBS units are representative of digital units in general, such units are likely to be useful in field research with clinical populations. The digital models are relatively easy to use, are inexpensive compared to ambulatory monitors, and possess acceptable levels of reliability and validity (at least when assessments are taken by a person experienced with the units). We do not know the extent that these results would be altered if subjects monitored their own blood pressures. Of the two digital instruments tested in the present study, the slightly better performance of the IBS compared to the Sears may be negligible considering differences in price: $700 versus $200. It appears that recent technological advancements in the production of portable blood pressure/pulse units have increased their reliability and made them more suitable for clinical research. This is particularly relevant in light of recent research indicating that home-based blood pressure readings, taken by patients themselves, are more reliable indicators of average blood pressure level than those readings taken in a physician's office (Kleinert et al., 1984). It is likely that these instruments will assume a more important role in the evaluation of hypertension both in clinical settings and possibly in field research.

REFERENCES

Arena, J. G., Blanchard, E. B., Andrasik, F., Cotch, P. A., & Myers, P. E. (1983). Reliability of psychophysiological assessment. *Behavior Research & Therapy*, **21**, 447-460.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychological Bulletin*, **56**, 81-105.

Cicchetti, D. V. (1972). Extension of multiple range tests to interaction tables in the analysis of variance: A rapid approximate solution. *Psychology Bulletin*, **77**, 405-408.

Devereux, R. B., Pickering, T. G., Harshfield, G. A., Kleinert, M. D., Denby, L., Clark, L., Pregibon, D., Jason, M., Kleiner, B., Borer, J. S., & Laragh, J. H. (1983). Left ventricular hypertrophy in patients with hypertension; importance of blood pressure response to regularly recurring stress. *Circulation*, **68**, 470-476.

Falkner, B., Kushner, H., Onesti, G., & Angelakos, E. T. (1981). Cardiovascular characteristics in adolescents who develop essential hypertension. *Hypertension*, **3**, 521-527.

Kleinert, H. D., Harshfield, G. A., Pickering, T. G., Devereux, R. B., Sullivan, R. A., Marion, R. M., Mallory, W. K., & Laragh, J. H. (1984). What is the value of home blood pressure measurement in patients with mild hypertension? *Hypertension*, **6**, 574-578.

Labarthe, D. R., Hawkins, C. M., & Remington, R. D. (1973). Evaluation of performance of selected devices for measuring blood pressure. *American Journal of Cardiology*, **32**, 546-553.

Light, K. C., & Obrist, P. A. (1983). Task difficulty, heart rate reactivity, and cardiovascular responses to an appetitive reaction time task. *Psychophysiology*, **20**, 301-312.

Obrist, P. A. (1981). *Cardiovascular psychophysiology: A perspective.* New York: Plenum.

Sher, K. J., Walitzer, K., Mannion, J., & Hammer, G. (1984, May). *The stability of stress responsiveness*. Paper presented at the Annual Convention of the Society of Behavioral Medicine, Philadelphia.

Steptoe, A. (1980). Blood pressure. In I. Martin & P. H. Venables (Eds.), *Techniques in psychophysiology* (pp. 247-274). New York: Wiley.

Williamson, D. A., Waters, W. F., Bernard, B., Faulstich, M., & Blouin, D. C. (1985, March). *Test-re-test reliability of psychophysiological assessment*. Paper presented at the Annual Meeting of the Society of Behavioral Medicine, New Orleans.

NOTES

1. Similar data collected during the mental arithmetic task are available upon written request to the first author.

2. Similar data collected during the Pac-Man task are available upon written request to the first author.

(Manuscript received February 26, 1985;
revision accepted for publication June 5, 1985.)