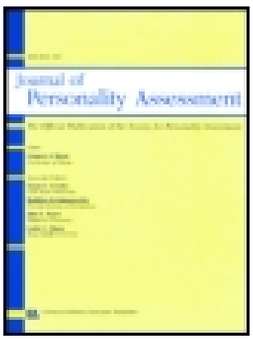

Routledge
Taylor & Francis Group

Journal of Personality Assessment

ISSN: 0022-3891 (Print) 1532-7752 (Online) Journal homepage: http://www.tandfonline.com/loi/hjpa20

Comparison of Beck Depression Inventories-IA and-II in Psychiatric Outpatients

Aaron T. Beck , Robert A. Steer , Roberta Ball & William F. Ranieri

To cite this article: Aaron T. Beck , Robert A. Steer , Roberta Ball & William F. Ranieri (1996) Comparison of Beck Depression Inventories-IA and-II in Psychiatric Outpatients, Journal of Personality Assessment, 67:3, 588-597, DOI: 10.1207/s15327752jpa6703_13

To link to this article: https://doi.org/10.1207/s15327752jpa6703_13

JOURNAL OF PERSONALITY ASSESSMENT, 1996, *67*(3), 588–597

ARTICLES

Comparison of Beck Depression Inventories –IA and –II in Psychiatric Outpatients

Aaron T. Beck
Beck Institute for Cognitive Therapy and Research
Bala Cynwyd, Pennsylvania

Robert A. Steer, Roberta Ball, and William F. Ranieri
University of Medicine and Dentistry of New Jersey
School of Osteopathic Medicine

The amended (revised) Beck Depression Inventory (BDI–IA; Beck & Steer, 1993b) and the Beck Depression Inventory–II (BDI–II; Beck, Steer, & Brown, 1996) were self-administered to 140 psychiatric outpatients with various psychiatric disorders. The coefficient alphas of the BDI–IA and the BDI–II were, respectively, .89 and .91. The mean rating for Sadness on the BDI–IA was higher than it was on the BDI–II, but the mean ratings for Past Failure, Self-Dislike, Change in Sleeping Pattern, and Change in Appetite were higher on the BDI–II than they were on the BDI–IA. The mean BDI–II total score was approximately 2 points higher than it was for the BDI–IA, and the outpatients also endorsed approximately one more symptom on the BDI–II than they did on the BDI–IA. The correlations of BDI–IA and BDI–II total scores with sex, ethnicity, age, the diagnosis of a mood disorder, and the Beck Anxiety Inventory (Beck & Steer, 1993a) were within 1 point of each other for the same variables.

The amended (revised) Beck Depression Inventory (BDI–IA; Beck & Steer, 1993a) is one of the most widely used measures for assessing the severity of depression in psychiatric patients and screening for possible depression in normal populations (Archer, Maruish, Imhof, & Piotrowski, 1991; Piotrowski & Keller, 1992; Piotrowski, Sherry, & Keller, 1985). Comprehensive reviews about the original Beck Depression Inventory (BDI–I; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) and the

BDI–IA have concluded that both scales have clinical utility and display reliable psychometric characteristics across a broad spectrum of both clinical and nonclinical populations (Beck, Steer, & Garbin, 1988; Steer, Beck, & Garrison, 1986). However, Moran and Lambert (1983) compared the BDI–IA's symptom content with the criteria for Affective Disorders as presented in the *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed. [*DSM–III];* American Psychiatric Association, 1980) and reported that the BDI–IA adequately met only six out of the nine *DSM–III* criteria. The BDI–IA requested information about decreases in appetite and sleep, but did not ask about increases in these two symptoms. Moran and Lambert (1983) also stressed that the BDI–IA did not include any items addressing psychomotor activity and agitation. Similar questions about the same symptom content in the BDI–I were raised by Vredenburg, Krames, and Flett (1985).

With the releases of the *DSM–III–R* and *DSM–IV* (American Psychiatric Association, 1987, 1994), Beck and his associates decided that there was a need to modify the BDI–IA to make it more consonant with *DSM–III–R/DSM–IV* criteria, and the Beck Depression Inventory–II[1] (BDI–II; Beck, Steer, & Brown, 1996) is the upgraded version of the BDI–IA. It was specifically constructed to measure the severity of self-deported depression in adolescents and adults according to *DSM–IV* criteria for diagnosing depressive disorders. Detailed descriptions about how the BDI–II was pilot-tested and normative data about its reliability and validity based on samples of 500 psychiatric outpatients and 120 undergraduates are presented by Beck et al. (1996). Briefly, items addressing *DSM–IV* depressive criteria that had not been included in the BDI–IA, such as Agitation, Concentration Difficulty, and Worthlessness, were written and pilot-tested along with BDI–IA items whose statement options were partially or completely reworded. The BDI–IA Weight Loss, Body Image Change, and Somatic Preoccupation items were dropped from the BDI–II and replaced with Concentration Difficulty, Worthlessness, and Loss of Energy items. Only three BDI–IA items were not reworded—Punishment Feelings, Suicidal Thoughts or Wishes, and Loss of Interest in Sex. Eighteen items addressing similar symptom domains in the BDI–IA were retained in the BDI–II.

The old BDI–IA labels for the symptom groups were also changed in some instances in the BDI–II. For example, the label, Self-Accusations, was changed to Self-Criticalness because the alternative statements in the BDI–II now better reflect the latter label than they do the former label. Additional clinical nuances were attached to some items. For example, Social Withdrawal, which only stresses interpersonal relationships in the BDI–IA, now emphasizes both people and activities in the BDI–II Loss of Interest item.

Furthermore, the directions for the BDI–II were changed from those given for the BDI–IA. The time frame for the BDI–II ratings was extended from 1 week in

[1]The Beck Depression Inventory–II is available from The Psychological Corporation, 555 Academic Court, San Antonio, TX 78204–2498.

the BDI–IA to 2 weeks to be consistent with *DSM–IV* criteria for major depression disorders. Respondents are now asked to describe themselves for the "Past Two Weeks, Including Today." Like the BDI–IA, the BDI–II is scored by summing the highest ratings for each of the 21 items. Each item is rated on a 4-point scale ranging from 0 to 3, and the total scores can range from 0 to 63. According to Beck et al. (1996), BDI–II total scores ranging from 0 to 13 represent "Minimal" depression; total scores from 14 to 19 are "Mild;" total scores from 20 to 28 are "Moderate;" and total scores from 29 to 63 are "Severe."

Beck et al. (1996) provided evidence that the BDI–II possesses adequate reliability and validity for clinical purposes. For example, they found that for 26 outpatients who completed the BDI–II before their first and second cognitive therapy sessions, the 1-week test–retest reliability was high ($r = .93, p < .001$). With respect to its convergent and discriminant validities, Beck, Steer, and Brown (1996) also reported that the BDI–II was more positively correlated with the revised Hamilton Psychiatric Rating Scale for Depression (Riskind, Beck, Brown, & Steer, 1987; $r = .71$) than it was with the revised Hamilton Rating Scale for Anxiety (Riskind et al., 1987; $r = .47$) in 87 outpatients. However, Beck et al. (1996) did not describe whether respondents' mean ratings for the 18 similar symptoms carried over from the BDI–IA to the BDI–II were comparable for both scales, nor do they indicate whether the correlations of the BDI–IA and the BDI–II total scores for the same respondents displayed comparable relationships with respect to psychosocial characteristics (such as sex, ethnicity, etc). The BDI–I and BDI–IA have repeatedly been found to have minimal ($r < .30$), usually nonsignificant correlations with sex, ethnicity, and age (Beck et al., 1988). Is the BDI–II also minimally correlated with these background characteristics?

The purposes of this study were (a) to compare the mean ratings of psychiatric outpatients for the 18 similar symptoms contained in the BDI–IA and BDI–II; and (b) to ascertain whether both instruments' total scores display comparable relationships with respect to sex, ethnicity, age, the diagnosis of a *DSM–IV* principal mood disorder, and overall severity of self-reported anxiety.

METHOD

Sample

The sample was composed of 140 psychiatric outpatients who were evaluated by the University of Medicine and Dentistry of New Jersey School of Osteopathic Medicine's Department of Psychiatry. There were 94 (67%) women and 46 (33%) men. The mean age was 37.60 ($SD = 16.84$) years old. There were 118 (84%) Whites, 4 (3%) Blacks, 1 (1%) Hispanic, and 17 (12%) Asians.

All of the patients were diagnosed by experienced psychiatrists according to *DSM–IV* criteria. However, although the psychiatrists were actively involved in

teaching *DSM–IV* diagnostic criteria to psychiatric residents, no interjudge agreement study was conducted with respect to the present diagnoses. There were 65 (46%) patients who were diagnosed with mood disorders; 23 (16%) with anxiety disorders; 30 (21%) with adjustment disorders; and 22 (16%) with various other types of disorders. For the 65 outpatients with mood disorders, there were 17 (12%) outpatients diagnosed with single-episode major-depressive disorders, and 15 (11%) diagnosed with recurrent-episode major-depressive disorders. Bipolar disorders were diagnosed for 10 (7%), dysthymic disorders for 21 (15%), and depressive disorders not-otherwise-specified (NOS) for 2 (1%).

Instrument

In addition to the BDI–II and the BDI–IA, the outpatients were also administered the Beck Anxiety Inventory (BAI; Beck & Steer, 1993a). The BAI is a 21-item self-report instrument for measuring the severity of anxiety that was specially developed to minimize its relationship with depression as measured by the BDI–IA. This instrument was specifically chosen for this study to ascertain if its correlation with the BDI–II would be comparable to that found for the BDI–IA. In using the BAI, patients rate the severity of each symptom on a 4-point scale ranging from 0 (*Not at all*) to 3 (*Severely—I could barely stand it*). A total score is calculated by summing the severity ratings for all 21 items. The BAI has been described by Beck and Steer (1993a) as displaying high concurrent validity with other self-report and clinical rating scales of anxiety (Beck & Steer, 1993a). In this study, it had high internal consistency ($\alpha = .90$).

Procedure

All of the outpatients completed the Beck scales as part of a standard intake battery of psychological tests. Patient participation was voluntary, and signed written consent forms were obtained before the instruments were administered. The order in which the outpatients were asked to complete the BDI–IA and the BDI–II was alternated, and the BAI was always the second instrument that was administered. Therefore, 70 outpatients completed the BDI–IA before being administered the BDI–II as the third instrument, and 70 outpatients were administered the BDI–II before being given the BDI–IA as the third scale.

RESULTS

Order of Administration

Because it usually takes less than 10 min to complete the BAI, the present outpatients were being administered the BDI–IA and the BDI–II within such a short

time interval that we had to assume that the outpatients would remember some of the items from both instruments. In fact, several outpatients complained about being asked "the same questions twice in a row." We thus had to ascertain whether the order in which the BDI–IA and the BDI–II had been administered affected the patients' responses to all three Beck instruments.

A one-way repeated-measures multivariate analysis of variance was performed to compare the BDI–IA, BDI–II, and BAI total scores by the order of administration of the BDI–II (1 = First, 2 = Third). The mean BDI–IA, BDI–II, and BAI total scores of the 70 outpatients who were administered the BDI–II first were, respectively, 20.20 (*SD* = 11.14), 22.17 (*SD* = 11.46), and 17.19 (*SD* = 9.87); and the mean BDI–IA, BDI–II, and BAI total scores of the 70 outpatients who completed the BDI–II third were, respectively, 20.50 (*SD* = 11.32), 22.56 (*SD* = 12.44), and 17.10 (*SD* = 12.16). The mean profiles of the total scores were comparable regardless of the order of administration of the BDI–II or BDI–IA, Wilks's lambda = .99, $F(3, 136) = .03$, *ns*. Therefore, order of administration was not controlled for in subsequent analyses.

Mean Comparisons

Table 1 presents the means and standard deviations of the 21 BDI–IA and the 21 BDI–II items along with the total scores and total number of different symptoms that were endorsed by the 140 outpatients for both instruments. The correlations between the 18 items measuring similar symptoms are also presented in Table 1 along with the corresponding paired *t* statistics that were used to test for mean differences between these ratings. The BDI–IA and BDI–II labels and item numbers are also listed in Table 1. A Bonferroni adjustment (alpha / 20) was used to control for the familywise error rate in determining whether a correlation or paired *t* test was significant. All of the positive correlations between the 18 similar symptoms shown in Table 1 were significant beyond the .001 level, one-tailed test. These correlations ranged from .38 for Irritability to .95 for Loss of Interest in Sex.

As Table 1 shows, the mean BDI–II total score was 2.01 points higher than that of the BDI–IA total score, paired $t(139) = 6.03$, $p < .001$. The correlation between the BDI–IA and BDI–II was .93, $p < .001$. The mean total scores for both instruments were within the moderately severe range of depression according to Beck and Steer's (1993b) and Beck et al.'s (1996) cutoff score guidelines. Table 1 also reveals that the outpatients had endorsed approximately one more item on the BDI–II than they had reported on the BDI–IA, paired $t(139) = 7.35$, $p < .001$.

With respect to the 18 similar symptoms measured by both the BDI–IA and the BDI–II, there were five significant mean differences (Table 1). The BDI–IA mean rating for Sadness was higher than it was for the BDI–II, but the mean ratings for Past Failure, Self-Dislike, Change in Sleeping Pattern, and Change in Appetite were

TABLE 1
Means, Standards Deviations, and Correlations Between Beck Depression Inventories–IA and –II

BDI–IA			*BDI–II*				
Item Label	*M*	*SD*	*Item Label*	*M*	*SD*	*r*	*t(139)*
Similar items							
1. Sadness	1.24	.98	1. Sadness	1.05	.95	.85	4.38**
2. Pessimism	1.01	.91	2. Pessimism	.95	.88	.74	1.05
3. Sense of failure	.76	.86	3. Past failure	.91	.94	.88	3.83*
4. Self-dissatisfaction	1.22	.93	4. Loss of pleasure	1.16	.87	.80	1.18
5. Guilt	.81	.91	5. Guilty feelings	.85	.91	.86	.87
6. Punishment	.93	1.21	6. Punishment feelings	.81	1.15	.82	1.90
7. Self-dislike	1.01	.93	7. Self-dislike	1.41	1.07	.67	5.77**
8. Self-accusations	1.14	.83	8. Self-criticalness	1.07	.92	.60	1.08
9. Suicidal ideas	.46	.66	9. Suicidal thoughts	.41	.56	.82	1.35
10. Crying	1.16	1.11	10. Crying	1.16	1.09	.84	0
11. Irritability	1.10	.82	17. Irritability	1.10	.88	.38	0
12. Social withdrawal	.87	.85	12. Loss of interest	1.05	.92	.70	1.10
13. Indecisiveness	1.19	1.00	13. Indecisiveness	1.19	1.04	.87	.16
15. Work difficulty	1.19	.91	15. Loss of energy	1.17	.86	.60	.21
16. Insomnia	1.33	.99	16. Change in sleeping	1.63	1.10	.56	3.80*
17. Fatigability	1.07	.93	20. Tiredness or fatigue	1.23	.96	.67	2.41
18. Loss of appetite	.67	.90	18. Change in appetite	.96	1.00	.59	4.03**
21. Loss of libido	.96	1.10	21. Loss of interest in sex	.96	1.11	.95	0
Different items							
14. Body image change	.92	1.05	11. Agitation	1.20	.87		
19. Weight loss	.51	.88	14. Worthlessness	.78	.94		
20. Somatic preoccupation	.78	.98	19. Concentration difficulty	1.30	.86		
Total score	20.35	11.19		22.36	11.92	.94	6.03**
Number of items endorsed	12.74	5.00		13.89	5.08	.93	7.35**

Note. $N = 140$. BDI–IA = Beck Depression Inventory–IA; BDI–II = Beck Depression Inventory–II. Bonferroni adjusted (alpha / 20).
*$p < .01$. **$p < .001$.

higher for the BDI–II than they were the BDI–IA. (The BDI–II labels are here used to describe the symptoms.) The mean differences with respect to the ratings for changes in appetite and sleep were expected because the BDI–II Change in Appetite and Change in Sleeping Pattern symptoms now, respectively, include statements about appetite and sleep increases, instead of asking just about appetite and sleep decreases. In fact, with respect to the BDI–II Change in Sleeping Pattern item, 28 (20%) outpatients reported that they were sleeping more, 92 (66%) indicated that they were sleeping less, and 20 (14%) described no change. For the BDI–II Change in Appetite, 62 (44%) described decreases in appetite, 20 (14%) reported increases in appetite, and 58 (41%) indicated no change in appetite. The rewording of these items had detected more potentially depressed outpatients than the old wording would have; 14% were hungrier, and 20% were sleepier.

Item Analyses

The coefficient alpha of the BDI–IA for the 140 outpatients was .89, and it was .91 for the BDI–II. These coefficients indicated that both instruments had comparably high internal consistencies. All 21 of the corrected item-total correlations for the BDI–IA were significant beyond the .05 level, one-tailed test, even after using a Bonferroni adjustment (alpha / 21) to control for the familywise error rate. The BDI–IA correlations ranged from .26 (Weight Loss) to .68 (Self-Dissatisfaction [anhedonia]). The lower range of the corrected-item correlations for the BDI–II was slightly higher, and all 21 of the corrected item-total correlations for the BDI–II were significant beyond the .001 level, one-tailed test, even after using a Bonferroni adjustment (alpha / 21). The lowest correlation for the BDI–II was .39 (Loss of Interest in Sex), and the highest correlation was .68 (Loss of Pleasure [anhedonia]). It is interesting to note that the highest corrected item-total correlation in both instruments is for anhedonia.

Correlations With Selected Variables

The point-biserial correlations of the BDI–IA and BDI–II total scores with sex (0 = Male, 1= Female), ethnicity (0 = Non-White, 1 = White), age (years), and the diagnosis of a *DSM–IV* principal mood disorder (0 = Nonmood, 1 = Mood) for the 140 outpatients are shown in Table 2. Neither instrument was significantly associated with the three background characteristics, and both were positively ($p < .001$) related to being diagnosed with a mood disorder, even after using a Bonferroni adjustment (alpha / 5). The positive correlations of the BAI total scores with the BDI–IA and the BDI–II total scores in Table 2 are highly significant ($p < .001$) and positive in magnitude. Furthermore, all of the correlations of the BDI–IA and

TABLE 2
Correlations of Beck Depression Inventories–IA and –II Total Scores With Sex, Ethnicity, Age, Diagnosis of a Principal Mood Disorder, and the Beck Anxiety Inventory

	BDI–IA	*BDI–II*
Variable	*r*	*r*
Sex (0 = Male, 1 = Female)	.16	.17
Ethnicity (0 = non-White, 1 = White)	.12	.13
Age (years)	–.05	–.06
Principal disorder (0 = nonmood, 1 = mood)	.24*	.23*
BAI total score	.65**	.66**

Note. $N = 140$. BDI–IA = Beck Depression Inventory–IA; BDI–II = Beck Depression Inventory–II; BAI = Beck Anxiety Inventory. Bonferroni adjusted (alpha / 5).
*$p < .01$. **$p < .001$.

BDI–II total scores with the same variables in Table 2 are within 1 point of each other and obviously comparable in terms of magnitude.

DISCUSSION

The results indicate that the BDI–IA ($\alpha = .89$) and BDI–II ($\alpha = .91$) have comparable high levels of internal consistency, and both instruments contain 21 symptoms that are all positively correlated with self-reported depression. The outpatients endorsed at least one more symptom on the BDI–II than they did on the BDI–IA, and the mean BDI–II total score was approximately 2 points higher than it was for the BDI–IA.

With respect to mean differences between the ratings of the 18 items that assess similar symptom domains in both instruments, the outpatients described more sadness on the BDI–IA than they did on the BDI–II. In contrast, the mean ratings for Past Failure, Self-Dislike, Change in Sleeping Pattern, and Change in Appetite were higher on the BDI–II than they were on the BDI–IA. Because the BDI–II includes statements about sleep and appetite increases as well as decreases, the latter two mean differences were expected.

This study was unable to test whether the BDI–II was a better discriminator of specific mood disorders (i.e., single-episode major depression vs. dysthymia) than the BDI–IA because the accuracy of the specific diagnoses was suspect; no structured interviews had been used by the present psychiatrists in establishing their diagnoses. The results only indicate that the BDI–IA and the BDI–II total scores were comparable in estimating being diagnosed with a *DSM–IV* mood disorder.

Obviously, future research needs to investigate how effectively the BDI–II, as opposed to the BDI–IA, differentiates specific types of *DSM–IV* mood disorders.

Over the past 35 years, a substantial database has accumulated about the clinical utility of the BDI–I and the BDI–IA (Beck, Steer, & Garbin, 1988; Steer, Beck, & Garrison, 1986), and clinicians and researchers may be wary about adopting a new instrument, such as the BDI–II, that has not been as extensively studied as the BDI–I and BDI–IA have been. However, based on the clinical cutoff score guidelines that have been proposed by Beck and Steer (1993b) for the BDI–II and Beck et al. (1996) for the BDI–IA, both scales indicated that the present outpatients were moderately depressed. The correlations of the BDI–IA and the BDI–II with sex, ethnicity, age, the diagnosis of a mood disorder, and the BAI were also comparable and fell within 1 point of each other for the same variables. Therefore, for all practical purposes, the BDI–IA and the BDI–II displayed a similar pattern of relationships with the same psychosocial characteristics. The transition from the usage of the BDI–IA to that of the BDI–II should introduce no meaningful interpretative problems.

With respect to future research about the relationships between the BDI–IA and the BDI–II, the comparability of the BDI–IA and the BDI–II with inpatients needs to be studied because inpatients should reflect patients with more severe levels of depression. Furthermore, because the present sample of outpatients was 84% White and drawn from a suburban, middle- to upper-middle class community near a large east-coast city, the psychometric characteristics of the BDI–II and its relationship with the BDI–IA must be evaluated in clinical samples representing higher proportions of minorities and different socioeconomic backgrounds.

REFERENCES

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Archer, R. P., Maruish, M., Imhof, E. A., & Piotrowski, C. (1991). Psychological test usage with adolescent clients: 1990 survey findings. *Professional Psychology: Research and Practice, 22,* 247–252.

Beck, A. T., & Steer, R. A. (1993a). *Manual for the Beck Anxiety Inventory.* San Antonio, TX: Psychological Corporation.

Beck, A. T., & Steer, R. A. (1993b). *Manual for the Beck Depression Inventory.* San Antonio, TX: Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory–II.* San Antonio, TX: Psychological Corporation.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8,* 77–100.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Moran, P. W., & Lambert, M. J. (1983). A review of current assessment tools for monitoring changes in depression. In M. S. Lambert, E. R. Christensen, & S. DeJulio (Eds.), *The assessment of psychotherapy outcome* (pp. 263–303). New York: Wiley.

Piotrowski, C., & Keller, J. W. (1992). Psychological testing in applied settings: A literature review from 1982–1992. *Journal of Training in the Practice of Professional Psychology, 6,* 74–82.

Piotrowski, C., Sherry, D., & Keller, J. W. (1985). Psychodiagnostic test usage: A survey of the Society for Personality Assessment. *Journal of Personality Assessment, 49,* 115–119.

Riskind, J. H., Beck, A. T., Brown, G., & Steer, R. A. (1987). Taking the measure of anxiety and depression: Validity of the reconstructed Hamilton scales. *Journal of Nervous and Mental Disease, 175,* 474–479.

Steer, R. A., Beck, A. T., & Garrison, B. (1986). Applications of the Beck Depression Inventory. In N. Sartorius & T. A. Ban (Eds.), *Assessment of depression* (pp. 121–142). New York: Springer-Verlag.

Vredenburg, K., Krames, L., & Flett, G. L. (1985). Reexamining the Beck Depression Inventory: The long and short of it. *Psychological Reports, 57,* 767–778.

Robert A. Steer
Department of Psychiatry
University of Medicine and Dentistry of New Jersey
School of Osteopathic Medicine
40 East Laurel Road
PCC Room 218
Stratford, NJ 08084–1350

Received January 19, 1996
Revised March 18, 1996