

---

# A Comparison of the Millon Behavioral Medical Diagnostic and Millon Behavioral Health Inventory With Medical Populations



Edward A. Wise

*Mental Health Resources, Memphis*



David L. Streiner

*University of Toronto*

There is a lack of normative data on broadband omnibus types of personality tests with medical populations. In fact, the only two tests normed on medical populations are the Millon Behavioral Medicine Diagnostic (MBMD) and the Millon Behavioral Health Inventory (MBHI). The internal consistency, test-retest reliabilities, and validity studies of these instruments are reviewed and compared in an effort to aid clinicians in discerning their relative psychometric strengths and weaknesses. Due to the lack of validity studies with the MBMD and the fact that reliability limits the ceiling of validity coefficients, the MBMD has yet to meet the challenges it was designed to meet. Implications for practice are addressed. © 2010 Wiley Periodicals, Inc. *J Clin Psychol* 66:1281–1291, 2010.

---

A psychological evaluation is recommended or required prior to numerous types of medical interventions, such as spinal surgery (Block, Gatchel, Deardorff, & Guyer, 2003), chronic pain interventions (Gatchel & Weisberg, 2000), organ transplantation (Collins & Labott, 2007), and bariatric surgery (LeMont, Moorehead, Parish, Redo, & Ritz, 2004). For example, a multidisciplinary treatment team is necessary to assist surgeons in evaluating prospective bariatric surgery patients (Santry, Chin, Cagney, Alverdy, & Lauderdale, 2006). Surveys have found that inclusion of mental health professionals in the assessment of bariatric surgery patients is the norm rather than the exception (Bauchowitz et al., 2005; Santry et al., 2006). Further, almost 90% of the programs responding to a survey indicated they required all patients to undergo

---

Correspondence concerning this article should be addressed to: Edward A. Wise, Mental Health Resources, PLLC, 1037 Cresthaven Road, Memphis, TN 38119; e-mail: EdWisePhD@gmail.com

psychological evaluations, and over 80% of bariatric surgery teams rated these evaluations as “very valuable” or “valuable” (Bauchowitz et al., 2005).

Sarwer and Wadden (2005) indicated that the mental health evaluations may help to identify comorbid mental health conditions in patients applying for the surgery. Ritz (2006) advocated that the evaluation should also identify psychosocial risk factors that may interfere with a patient reaching optimal medical goals, and making the psychosocial, emotional, interpersonal, and behavioral adjustments that await them postsurgically.

Psychologists completing presurgical evaluations typically use traditional psychological tests and a clinical interview in their work with bariatric patients. Walfish, Vance, and Fabricatore (2007) found that 85% of psychologists surveyed used a clinical interview and psychological tests to conduct these evaluations. The results from Walfish et al. (2007), Bauchowitz et al. (2005), and Fabricatore, Crerand, Wadden, Sarwer, and Krasucki (2006) indicate that the most frequently used objective broadband personality tests in bariatric presurgical evaluations are the Minnesota Multiphasic Personality Inventory-II (MMPI-II; Butcher et al., 2001), the Personality Assessment Inventory (PAI; Morey, 2007), the Millon Clinical Multiaxial Inventory-III (MCMI-III; Millon, Davis, & Millon, 1997), the Millon Behavioral Medicine Diagnostic (MBMD; Millon, Antoni, Millon, Minor, & Grossman, 2006), and the Millon Behavioral Health Inventory (MBHI; Millon, Green, & Meagher, 1982). Only the latter two tests, however, have been extensively normed on diverse groups of medical patients. Although both the MMPI-II and PAI have been used with specific medical populations such as bariatric, back pain, and other patients, normative data for all of the scales across diverse medical populations are not readily available nor are there scoring reports for either of these instruments. The absence of scoring reports with medical norms makes it cumbersome for clinicians to compare their MMPI-2 or PAI scoring reports with those of similar medical patients, unless clinicians routinely refer to the literature to obtain reference norms on groups similar to the patient being evaluated. On the other hand, the MBHI provides normative data ( $N = 1,194$ ) obtained from surgery centers, pain clinics, dialysis units, and cancer programs, while the MBMD is comprised of normative data ( $N = 720$ ) obtained from patients being treated for cancer; diabetes; human immunodeficiency virus (HIV); and neurological, cardiac, pain, and other medical illnesses. Both scoring printouts are based on these norms.

In light of the fact that normative psychometric properties, such as reliability, are sample dependent, to use a test with norms derived from a community sample (i.e., MMPI-2 and PAI), one would have to assume that the community sample is representative of the medical population. For example, Keller and Butcher (1991) reviewed the MMPI-2 literature on pain patients and concluded “It will be vitally important to collect data on the new MMPI-2 items and scales in a general medical population to provide an appropriate comparison group... (and) to make more clinically relevant decisions such as differentiating chronic pain patients from a general medical population rather than from the normative and psychiatric groups...” (p. 223). Indeed, they cite numerous examples of prototypical high points and code types (e.g., neurotic triad, conversion V, etc.) obtained with pain patients where traditional interpretations would be off the mark, stating that item-level analyses are often required to clarify the meaning of these code types.

In fact, Swenson, Pearson, and Osborne (1973) demonstrated significant differences between 50,000 Mayo clinic medical outpatients and the original MMPI norms. Replicating this work, Colligan et al. (2008) reported on 1,243 Mayo clinic

patients and demonstrated “numerous and statistically significant differences between the normative reference values (mean, 50; *SD*, 10) for the MMPI-2 validity and clinical scales presented in the MMPI-2 manual and those obtained from our contemporary Midwestern medical outpatients (for both women and men)” (p. 105). In fact, the discrepancies were so compelling as to result in Colligan et al. (2008) developing new normative tables for medical outpatients “to allow comparability of scale elevations, or departure from the mean, in terms of percentile ranks, across all 18 standard validity and clinical scales of the MMPI-2...” (p. 105).

Similarly, the same logic would apply to bariatric surgical candidates, who are increasingly being referred for psychological evaluations (Bauchowitz et al., 2005; Santry et al., 2006). Obese medical patients preparing for bariatric surgery should optimally not be compared to the normative community group, but to a group of medical patients, or possibly even obese presurgical candidates. Obese and bariatric populations tend to have higher prevalence rates of depression and binge eating compared to the general population (Black, Goldstein, & Mason, 1992; de Zwaan, 2001; Dixon, Dixon, & O’Brien, 2003; Kalarchian, Wilson, Brolin, & Bradley, 1998; Roberts, Kaplan, Shema, & Strawbridge, 2000). The higher prevalence rates of these psychiatric disorders themselves would justify the use of tests, whose norms are based on clinical, or medical, but not community, samples. This is because when prevalence rates are high, positive results are more likely to be true. But this does not address the problem of how severe the depression or bingeing may be compared to other medical or bariatric patients. If one is interested in how this particular patient will fare compared to similar patients, a relevant normative group is required.

If there is reason to suspect that a particular individual is likely to be from a psychopathological clinical population, any of the above-mentioned tests are likely appropriate. Even then, however, the differential effects of having a psychopathological condition versus a medical condition could affect the test results and their interpretations. As an example, consider a male with a *T*-score of 70 on the MMPI-2 Depression scale from a psychotherapy outpatient setting, compared to an outpatient pain patient, who both endorse the following items: (10) I am about as able to work as I ever was; (18) I am troubled by attacks of nausea and vomiting; (31) I find it hard to keep my mind on a task or job; (33) I seldom worry about my health; (39) My sleep is fitful and disturbed; (45) I am in just as good physical health as most of my friends; (109) I seem to be as capable and smart as others around me; (233) I have difficulty starting to do things. If, upon interview, it is determined that these eight items legitimately reflect pain-related functional impairments and medication side effects, a traditional interpretation of this scale and *T*-score would not be appropriate. Furthermore, if one then eliminates these eight items for the pain patient protocol, the corresponding *T*-score is reduced to 57, indicating that this individual is not depressed. In addition to item analyses, keeping with the current example, the use of the Colligan et al. (2008) medical norms might clarify this picture, or the use of additional measures, such as the Beck Depression Inventory for Primary Care (Beck, Guth, Steer, & Ball, 1997), which eliminated contaminating somatic items and has 82% sensitivity and specificity, might also be administered to further clarify the presence of depression as both have medical patient norms.

To date, however, there are only two objective, broadband personality tests normed on medical patients that routinely provide clinicians with medical patient norms: the MBMD and the MBHI. It is of interest to also note that both tests were constructed by the same test author and reflect Millon’s theory of personality (Millon, Millon, Meagher, Grossman, & Ramnath, 2004). This fact is important in

that the same theory of personality underlies both tests, which should increase their convergence.

The primary purpose of the MBHI test is to provide information to clinicians, including psychologists, physicians and nurses, who deal with physically ill and behavioral medicine patients. It can aid in the psychological assessment of these patients and facilitate the steps required to formulate a comprehensive treatment plan. It provides information regarding a patient's likely style of relating to health-care personnel, problematic psychosocial attitudes and stressors, as well as an individual's similarity to patients with psychosomatic complications or poor responses to either illness or treatment interventions (Millon, 2006).

The MBHI was developed using a "mixed clinical and non-clinical population... [which] led to a loss of diagnostic behavioral precision... The decision to develop a pure medical normative group, with appropriate reference norms, served as a major impetus for constructing the MBMD..." (Millon et al., 2006, p. 1). However, the MBHI was normed on "2,113 patients... in a number of medical settings, and 452 subjects... from non-clinical settings... (Millon et al., 1982, p. 10), whereas the MBMD was normed on a final patient sample of 720 (Millon et al., 2006, p. 24). Additionally, several of the MBHI scales that "were created entirely on empirical grounds (e.g., Gastrointestinal Susceptibility) have not fared (well)... Other scales with minor problems were renamed and reformulated, some with slightly different areas of focus (e.g., Future Despair became Future Pessimism). Most importantly, we added scales that are related to psychiatric status, stress moderators, and treatment prognostics" (Millon et al., 2006, p. 3).

The MBMD is composed of 165 items; it has 29 clinical scales, three response pattern scales, one validity indicator, and six negative health habits indicators. The response pattern scales and validity indicator are similar to those found in the other Millon instruments and the negative health habits indicators "cover lifestyle behaviors that have been shown to have the strongest contribution to the widest variety of health outcomes: Alcohol, Drug, Eating, Caffeine, Inactivity, and Smoking" (Millon et al., 2006, p. 12). Additionally, there are 11 coping style scales, which are derived from Millon's personality theory and that correspond to most of the personality disorders according to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)*; American Psychiatric Association, 1994). The MBMD was designed to "Identify patients who may have significant psychiatric problems and recommend specific interventions; pinpoint personal and social assets that may facilitate adjustment to physical limitations or lifestyle changes; determine whether patients need more communication and support in order to comply with prescribed medical regimens; and structure post-treatment plans and self-care responsibilities in the context of the patient's social network" (Pearson Assessments, 2009).

Although the MBMD has many similarities with MBHI, and is intended to replace it, we could not locate any published psychometric comparisons between them. However, there is enough similarity in the construction of these tests and their predecessors to indicate concerns about their psychometric properties. For example, as with the previous Millon scales, item overlap has been shown to be a significant problem (e.g., Craig, 1999). Similarly, the MBMD contains 165 items, measuring 38 scales, while the MBHI contains 150 items, assessing 21 scales. Such a high degree of item overlap has been shown to artificially increase the convergent validity correlations between scales,

while decreasing their discriminant validity, and produces factorial invariance. To counter the problem of item overlap, the test authors devised a system in which each item was weighted according to the extent to which it was deemed to be a prototype for each respective scale. However, Streiner, Goldberg, and Miller (1993) were among the first to demonstrate that item weights failed to increase the reliability or validity of the scales. Grove and Vrieze (2009) have recently criticized the Millon scales for relying on base rate (BR) transformation scores that adjust each scale to account for the probability of occurrence of various clinical and personality syndromes (i.e., the prevalence rates). They provide a detailed analysis reporting that the information needed to analyze “the raw score distributions, the tallied clinician diagnosis base rates, the epidemiological study-reported base rates, and the subjectively adjusted base rates” (p. 62) are not available for review and go so far as to state the tests fail to meet the minimal criteria specified by the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education’s (1999) *Standards for Educational and Psychological Testing*. They also point out large BR discrepancies between the instrument and national epidemiological data, which may have considerable consequences. Due to the similarity in test construction, the MBHI and MBMD may be expected to have similar psychometric deficiencies.

### Method

With this history of problems related to the Millon tests, it appears prudent to compare the internal consistency and test retest-reliabilities of the newest Millon inventory, the MBMD, with its predecessor, the MBHI, using data obtained from their respective test manuals. Similarly, we reviewed validity studies on both instruments obtained by searching PubMed and PsycINFO. We included coefficient alpha because it is a minimum requirement for internal consistency and it defines the ceiling for any validity coefficients. Alpha coefficients reflect a minimum standard that indicates the scale items are likely consistent and unacceptably low alpha’s indicate a clear lack of consistency. On the other hand, test-retest information is helpful to clinicians who are concerned about the reproducibility of the scores. We understand that different reliability coefficients might be of greater interest, depending on the sources of variance of interest, but these two reliability coefficients are also those that are consistently reported in most test manuals.

### Results

#### *Internal Consistency*

Although the MBMD contains Millon’s updated Coping Styles scales, which include the Dejected, Non-Conforming, and Denigrated scales, it is evident from Table 1 that the Coping Styles scales from the original MBHI perform somewhat better than those on the MBMD. More specifically, the two tests have comparable internal consistency coefficients for the Introversive, Inhibited, and Cooperative scales, while the MBHI shows higher internal consistency coefficients for the Sociable, Confident, Forceful, Respectful, and Sensitive scales. Clearly the MBMD contains parallel forms of the eight MBHI Coping Style scales.

An additional seven of the remaining 12 MBHI scales appear to have parallel forms on the MBMD. Based on the descriptions of the scales from their respective manuals, the following seven MBHI scales are followed by their comparable MBMD scales and corresponding alpha coefficients: Chronic Tension (.77)–Anxiety Tension (.82),

Table 1

*Millon Behavioral Medicine Diagnostic (MBMD) and the Millon Behavioral Health Inventory (MBHI) Reliability Coefficients*

MBMD	Internal consistency (Cronbach's $\alpha$ )	Test-retest reliability	MBHI	Internal consistency KR-20	Test-retest reliability
Coping Styles			Coping Styles		
Introversive	0.74	0.71	Introversive	0.72	0.79
Inhibited	0.85	0.82	Inhibited	0.84	0.84
Dejected	0.81	0.83			
Cooperative	0.68	0.87	Cooperative	0.68	0.81
Sociable	0.54	0.79	Sociable	0.82	0.83
Confident	0.61	0.83	Confident	0.66	0.86
Non-Conforming	0.67	0.80	Forceful	0.72	0.77
Forceful	0.67	0.83			
Respectful	0.62	0.90	Respectful	0.74	0.78
Oppositional	0.83	0.78	Sensitive	0.86	0.88
Denigrated	0.79	0.86			
Psychiatric Indications			Psychogenic Attitude		
Anxiety-Tension	0.82	0.86	Chronic Tension	0.77	0.90
Depression	0.89	0.80	Recent Stress	0.74	0.87
Cognitive Dysfunction	0.83	0.79	Premorbid Pessimism	0.90	0.85
Emotional Lability	0.81	0.83	Future Despair	0.86	0.78
Guardedness	0.76	0.88	Social Alienation	0.84	0.85
			Somatic Anxiety	0.86	0.79
Stress Moderators			Psychosomatic Correlate		
Illness Apprehension	0.86	0.85	Allergic Inclination	0.81	0.83
Functional Deficits	0.85	0.87	Gastrointestinal Sus	0.83	0.81
Pain Sensitivity	0.89	0.86	Cardiovascular Tend	0.85	0.79
Social Isolation	0.86	0.78			
Future Pessimism	0.87	0.79	Prognostic Index		
Spiritual Absence	0.88	0.92	Pain Treatment Resp	0.86	0.82
			Life Threat Reactivity	0.83	0.76
Treatment Prognostics			Emotional Vulnerability	0.82	0.59
Interventional Fragility	0.80	0.88			
Medication Abuse	0.65	0.72			
Information Discomfort	0.47	0.81			
Utilization Excess	0.76	0.83			
Problematic Compliance	0.62	0.84			
Management Guides					
Adjustment Difficulties	0.77	0.81			
Psych Referral	0.79	0.78			

Premorbid Pessimism (.90)–Dejected (.81), Future Despair (.86)–Future Pessimism (.87), Social Alienation (.84)–Social Isolation (.86), Somatic Anxiety (.86)–Illness Apprehension (.86), Pain Treatment Response (.86)–Pain Sensitivity (.89), and Emotional Vulnerability (.82)–Interventional Fragility (.80). Of the seven sets of scales, five obtained comparable alpha coefficients and only two sets showed a difference of .05 or more (Chronic Tension–Anxiety Tension and Premorbid Pessimism–Dejected). Four of the MBHI scales (Allergic Inclination, Gastrointestinal Susceptibility, Cardiovascular Tendencies, and Life Threat Reactivity) were abandoned because they “have not fared as well as... hoped” (Millon et al., 2006, p. 3).

Of the remaining eight MBMD scales, three of the Treatment Prognostic scales (Medication Abuse, Information Discomfort, and Problematic Compliance)

demonstrated internal reliability coefficients  $< .70$ . The other five scales (Functional Deficits, Spiritual Absence, Utilization Excess, Adjustment Difficulties, and Psych Referral) demonstrate more reasonable alphas, although all are below the level of  $.90$ , recommended by Nunnally and Bernstein (1994) as the lower limit for clinical scales.

### *Test-Retest Reliability*

The test-retest reliability coefficients for both the MBHI and the MBMD are all above  $.70$ , with the exception of the MBHI Emotional Vulnerability scale. It is of interest to note, however, that the MBMD test-retest data do not include the bariatric candidates and the testing occurred over 7 to 30 days, whereas the MBHI data was on “a general population” (Millon, 1982, p. 25) and their retest interval averaged 4.5 months. The MBHI demonstrates somewhat higher test-retest reliability for six of the eight Coping Style scales it measures.

### *Validity*

With a longer history, the MBHI has some published validity studies. For example, Brandwin, Trask, Schwartz, and Clifford (2000), Coffman and Brandwin (1999), and Harper, Chacko, Kotik-Harper, Young, and Gotto (1998) have demonstrated that MBHI predicts mortality in cardiac transplant patients, compliance, coping, and survival time. Pereira et al. (2004) demonstrated that the MBHI inhibited Coping Style scale predicted nonadherence to follow up visits with HIV-positive women, whereas Goodkin et al. (1992) reported that the MBHI active Coping Style scales were positively associated with natural killer cell cytotoxicity HIV-positive homosexual men. Lantinga, Krauss, Clark, and Schell (1988) showed the MBHI to be effective in discriminating psychogenic, organic, and mixed etiologies of erectile dysfunction. Bockian, Meager, and Millon (2000) reported that the MBHI was able to predict positive outcomes in pain rehabilitation patients (e.g., time sitting, standing, treadmill, hand grip strength, etc.), pain and headache patients (e.g., decreased pain, decreased medication, increased activity, etc.), cancer patients tolerance for uncomfortable medical procedures (hyperthermia), etc. On the other hand, Rowe, Downey, Faust, and Horn (2000) demonstrated that the MCMII-III outperformed the MBHI in predicting weight loss following gastropasty, while Herron, Turner, Ersek, and Weiner (1992) showed that the MBHI scales were not associated with lumbar laminectomy outcome.

In the 9 years since its introduction, we were only able to locate two validity studies related to the MBMD. Cipher, Kurain, Fulda, Snider, and Van Beest, 2007 reported that coping style groups predicted different trajectory of change curves for patients in physical rehabilitation. The other study (Cruss, Minor, Antoni, & Millon, 2007) demonstrated that the Medication Abuse scale predicted medication adherence in an HIV-positive sample. It is interesting to note, however, that these authors, the latter three of whom also authored the MBMD manual, used raw scores, not adjusted prevalence scores, on which to base their findings. They stated “...because prevalence scores are not linear, the raw scores from the MBMD, which are linear, were employed... (p. 280). Thus, it is questionable how generalizable the results are, as they scored the test in a nonstandardized manner that does not reflect the prevalence score transformations or prevalence score adjustments that would be applied in the routine automated scoring obtained in clinical practice.

### Discussion

The Coping Style scales assess personality characteristics “that reflect the cognitive, behavioral, and interpersonal strategies” patients generally employ to obtain pleasure or avoid pain (Millon et al., 2006, p. 15). Because these personality styles influence nearly all aspects of an individual’s behavior, it is important that the scales hypothesized to measure these more enduring characteristics demonstrate relatively high internal and test-retest reliability. Although both sets of Coping Style scales achieve reasonable test-retest reliability, the MBHI scales demonstrate comparatively higher internal consistency. Thus, it is less clear what the MBMD Cooperative, Sociable, Confident, Non-Conforming, Forceful, and Respectful scales are measuring. Although it is conceivable that these MBMD scales are capturing a more complex array of prototypical characteristics, it is also conceivable that they have become more diffuse and less well defined. In either event, the Coping Style parallel scales do not perform comparably from a reliability standpoint. Additionally, reliability sets the upper limit for validity coefficients; therefore, this finding has potential implications for the validity of the respective scales. It should be noted that the higher values of alpha found with the MBHI Coping Style scales are likely due to the fact that they are comprised of more than twice as many items ( $M = 37$ ; range = 32–48) as the MBMD Coping Style scales ( $M = 15$ ; range = 9–22). Similarly, three of the five Treatment Prognostics scales (Medication Abuse, Information Discomfort, and Problematic Compliance) obtained relatively low alpha coefficients and will likely prove somewhat difficult to interpret.

Five scales (Functional Deficits, Spiritual Absence, Utilization Excess, Adjustment Difficulties, and Psych Referral) demonstrate reasonable alpha and test-retest coefficients and appear to represent improvements from the MBHI to the MBMD. Whether or not these scales are the most parsimonious methods to assess functional impairments, spiritual beliefs, risk of treatment complications due to coping style, or need for a psychological or psychiatric referral remains an empirical question. The test authors acknowledge that their Utilization Excess scale may also be tapping into “genuine medical complications... feigning or simulation” (Millon et al., 2006, p. 39).

It is clear that five of the MBMD Coping Style scales are less internally consistent than the MBHI Coping Style scales. This sacrifice is significant in light of the personologic interpretive framework the Coping Style scales provide. Furthermore, the MBMD Coping Style scales with scores  $\geq 60$  are adjusted up or down in an attempt to “correct for over- or underreporting” (Millon et al., 2006, p. 74). Similarly, on the MBHI, the psychogenic attitude scale scores were adjusted based on some of the Coping Style scales (2, 4, 7, or 8).

The use of the Coping Style scales to alter other scale scores makes their psychometric properties particularly important. As demonstrated by Streiner et al. (1993) and recently detailed by Grove and Vrieze (2009), the evidence to support the exact number of points scales were adjusted up or down was not made explicit and no studies were provided that compared the classification efficiencies of adjusted versus unadjusted scores. Nonetheless, it is clear that adjustments based on scales with less than optimal reliabilities are utilized in such a manner to affect other scores, which at best results in an unknown effect on reliabilities and at worst compounds the problem with scale reliabilities. Streiner et al. (1993) as well as Grove and Vrieze (2009) have called for the test authors to either eliminate scale score adjustments or to provide the missing data (i.e., descriptive statistics, raw score distributions, tallied clinician diagnosis base rates, epidemiological study-reported base rates, and the subjectively adjusted base rates) to review and determine their accuracy.



Compared to the MBMD, the MBHI test-retest correlations ( $r$ ) were obtained over a longer period of time and the Coping Style scales demonstrated slightly higher  $r$ s. According to Millon's theory, "Coping styles are hypothesized to affect a wide range of medical outcomes... They may also be useful for predicting the best match between a patient's coping style and a specific type and form of psychosocial intervention... The scales... correspond to contemporary DSM personality disorders..." (Millon et al., 2006, p. 15). Thus, these characteristics are assumed to reflect enduring and pervasive characteristics that typify an individual's relationship with the world and may be expected to affect their relationship with medical providers, illness, other caregivers, etc. If such personality issues are of interest, the MBHI appears to have advantages from the standpoint of psychometric reliability.

Although we applaud the development of new tests normed on medical patients, we encourage practitioners to proceed cautiously in adopting new instruments, based on test construction methods that have proven to be problematic. With this in mind, it is also clear that the MBMD normative data demonstrate some scale reliability problems that will make clinical interpretations even more difficult, at least until other validity studies have been conducted. We recognize that the absence of validity comparisons is a significant limitation of this article, but this limitation reflects the state of the literature and strengthens the rationale of the need for such a review to aid practitioners in determining which instrument to rely upon. In the absence of additional validity data, it seems clear that scales with alphas  $<.70$  should be interpreted cautiously and that convergent validity should be sought for interpretive statements derived from these scales. We believe that the MBHI, which has a history of research, is also based on similar problematic methods of test construction, but because it has received greater research, continues to have a role in clinical practice, at least until such time that the MBMD has an established research base that demonstrates incremental validity beyond the MBHI. We agree with Grove and Vrieze (2009) that the underlying psychometric data, including descriptive statistics, raw score distributions, tallied clinician diagnosis base rates, epidemiological study-reported base rates, and the subjectively adjusted base rates for all of the Millon instruments, should be revealed. This will facilitate research on these tests and allow for better comparisons between the MBMD, MBHI, and MCMI-III to determine their relative strengths and weaknesses in general, and with medical populations in particular. More meaningful and relevant validity work with these instruments would advance practice with medical patients.

### References

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Bauchowitz, A.U., Gonderer-Frederick, L.A., Olsbrich, M.E., Azarbad, L., Ryee, M.Y., Woodson, M., et al. (2005). Psychosocial evaluation of bariatric surgery candidates: A survey of present practices. *Psychosomatic Medicine*, 67, 825–832.
- Beck, A.T., Guth, D., Steer, R.A., & Ball, R. (1997). Screening for major depression disorders in medical inpatients with the Beck Depression Inventory for Primary Care. *Behavior Research and Therapy*, 35, 785–791.
- Black, D.W., Goldstein, R.B., & Mason, E.E. (1992). Prevalence of mental disorder in 88 morbidly obese bariatric clinic patients. *American Journal of Psychiatry*, 149, 227–234.

- Block, A.R., Gatchel, R., Dearthoff, W., & Guyer, R. (2003). *The psychology of spine surgery*. Washington, DC: American Psychological Association.
- Bockian, N., Meager, S., & Millon, T. (2000). Assessing personality with the Millon Behavioral Health Inventory, the Millon Behavioral Medicine Diagnostic, and the Millon Clinical Multiaxial Inventory. In R.J. Gatchel & J.N. Weisberg (Eds.), *Personality characteristics of patients with pain*. Washington, DC: American Psychological Association.
- Brandwin, M., Trask, P., Schwartz, S., & Clifford, M. (2000). Personality predictors of mortality in cardiac transplant candidates and recipients. *Journal of Psychosomatic Research*, 49, 141–147.
- Butcher, J., Graham, J., Ben-Porath, Y., Tellegen, A., Dahlstrom, W., & Kaemmer, B. (2001). *MMPI-2 manual for administration, scoring and interpretation* (Rev. ed.). Minneapolis: University of Minnesota Press.
- Cipher, D.J., Kurian, A.K., Fulda, K.G., Snider, R., & Van Beest, J. (2007). Using the Millon Behavioral Medicine Diagnostic to delineate treatment outcomes in rehabilitation. *Journal of Clinical Psychology in Medical Settings*, 14, 102–112.
- Coffman, K., & Brandwin, M. (1999). The Millon Behavioral Health Inventory Life Threat Reactivity Scale as a predictor of mortality in patients awaiting heart transplantation. *Psychosomatics*, 40, 44–49.
- Colligan, R.C., Rasmussen, N.H., Agerter, D.C., Offord, K.P., Malinchoc, M., O'Byrne, M.M., et al. (2008). The MMPI-2: A Contemporary Normative Study of Midwestern Family Medicine Outpatients. *Journal of Clinical Psychology in Medical Settings*, 15, 98–119.
- Collins, C.A., & Labott, S. (2007). Psychological assessment of candidates for solid organ transplantation. *Professional Psychology: Research and Practice*, 38, 150–157.
- Craig, R.J. (1999). Overview and current status of the Millon Clinical Multiaxial Inventory. *Journal of Personality Assessment*, 72, 390–406.
- Cruess, D.G., Minor, S., Antoni, M., & Millon, T. (2007). Utility of the Millon Behavioral Medicine Diagnostic (MBMD) to predict adherence to highly active antiretroviral therapy (HAART) medication regimens among HIV-positive men and women. *Journal of Personality Assessment*, 89, 277–290.
- de Zwaan, M. (2001). Binge eating disorder and obesity. *International Journal of Obesity*, 25, S51–S55.
- Dixon, J.B., Dixon, M.E., & O'Brien, P.E. (2003). Depression in association with severe obesity. *Archives of Internal Medicine*, 163, 2058–2065.
- Fabricatore, A., Czerand, C., Wadden, T., Sarwer, D., & Krasucki, J. (2006). How do mental health professionals evaluate candidates for bariatric surgery? Survey results. *Obesity Surgery*, 16, 567–573.
- Gatchel, R.J., & Weisberg, J.N. (2000). *Personality characteristics of patients with pain*. Washington, DC: American Psychological Association.
- Goodkin, K., Blaney, N.T., Feaster, D., Fletcher, M.A., Baum, M.K., Mantero-Atienza, E., et al. (1992). Active coping style is associated with natural killer cell cytotoxicity in asymptomatic HIV-1 seropositive homosexual men. *Journal of Psychosomatic Research*, 36, 635–650.
- Grove, W.M., & Vrieze, S.I. (2009). An exploration of the base rate scores of the Millon Clinical Multiaxial Inventory–III. *Psychological Assessment*, 21, 57–67.
- Harper, R., Chacko, R., Kotik-Harper, D., Young, J., & Gotto, J. (1998). Self-report evaluation of health behavior, stress vulnerability, and medical outcome of heart transplant recipients. *Psychosomatic Medicine*, 60, 563–569.
- Herron, L., Turner, J.A., Ersek, M., & Weiner, P. (1992). Does the Millon Behavioral Health Inventory (MBHI) predict lumbar laminectomy outcome? A comparison with the Minnesota Multiphasic Personality Inventory (MMPI). *Journal of Spinal Disorders*, 5, 188–192.
- Kalarchian, M., Wilson, G.T., Brolin, R.E., & Bradley L. (1998). Binge eating in bariatric surgery patients. *International Journal of Eating Disorders*, 23, 89–92.

- Keller, L.S., & Butcher, J.N. (1991). *Assessment of chronic pain patients with the MMPI-2*. Minneapolis: University of Minnesota Press.
- Lantinga, L.J., Krauss, D.J., Clark, M.M., & Schell, R.E. (1988). Differentiating the etiology of male erectile dysfunction using the Millon Behavioral Health Inventory and self-report measures. *Journal of Sex and Marital Therapy*, 14, 263–268.
- LeMont, D., Moorehead, M., Parish, M., Redo, S., & Ritz, S. (2004). Suggestions for the pre-surgical psychological assessment of bariatric surgery candidates. Gainesville, FL: American Society for Bariatric Surgery.
- Millon, T. (2006). The Institute for Advanced Studies in Personology and Psychopathology. Retrieved March 3, 2010, from <http://www.millon.net/instruments/MBHI.htm>
- Millon, T., Antoni, M., Millon, C., Minor, S., & Grossman, S. (2006). *Millon Behavioral Medicine Diagnostic*. Minneapolis, MN: National Computer Systems.
- Millon, T., Davis, R., & Millon, C. (1997). *Millon Clinical Multiaxial Inventory-III* (2nd ed). Minneapolis, MN: National Computer Systems.
- Millon, T., Green, C., & Meagher, R. (1982). *Millon Behavioral Health Inventory manual* (3rd ed.). Minneapolis, MN: National Computer Systems.
- Millon, T., Millon, C., Meagher, S., Grossman, S., & Ramnath, R. (2004). *Personality disorders in modern life* (2nd ed.). New York: Wiley.
- Morey, L.C. (2007). *Personality Assessment Inventory* (2nd ed.). Odessa, FL: Psychological Assessment Resources, Inc.
- Nunnally, J.C., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Pearson Assessments. (2009). *Millon Behavioral Medicine Diagnostic*, product summary. Retrieved March 3, 2010, from <http://psychcorp.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid = PAg503&Mode = summary>
- Pereira, D., Antoni, M., Danielson, A., Simon, T., Efantis-Potter, J., & O'Sullivan, M. (2004). Inhibited interpersonal coping style predicts poorer adherence to scheduled clinic visits in human immunodeficiency virus infected women at risk for cervical cancer. *Annals of Behavioral Medicine*, 28, 195–202.
- Ritz, S. (2006). The bariatric psychological evaluation: A heuristic for determining the suitability of the morbidly obese patient for weight loss surgery. *Bariatric Nursing and Patient Care*, 1, 97–105.
- Roberts, R.E., Kaplan, G.A., Shema, S.J., & Strawbridge, W.J. (2000). Are the obese at greater risk for depression? *American Journal of Epidemiology*, 152, 163–170.
- Rowe, J., Downey, J., Faust, M., & Horn, M. (2000). Psychological and demographic predictors of successful weight loss following silastic ring vertical stapled gastroplasty. *Psychological Reports*, 86, 1028–1036.
- Santry, H.P., Chin, M.C., Cagney, K.A., Alverdy, J.C., & Lauderdale, D.S. (2006). The use of multidisciplinary teams to evaluate bariatric surgery patients: Results from a national survey in the USA. *Obesity Surgery*, 16, 59–66.
- Sarwer, D., & Wadden, T. (2005). The value of preoperative psychological evaluations of bariatric surgery candidates. *Surgery for Obesity and Related Diseases*, 1, 528–529.
- Streiner, D.L., Goldberg, J., & Miller, H.R. (1993). MCMI-II item weights: Their lack of effectiveness. *Journal of Personality Assessment*, 60, 471–476.
- Swenson, W.M., Pearson, J.S., & Osborne, D. (1973). *An MMPI source book: Basic item, scale, and pattern data on 50,000 medical patients*. Minneapolis, MN: University of Minnesota Press.
- Walfish, S., Vance, D., & Fabricatore, A. (2007). Psychological evaluation of bariatric surgery applicants: Procedures and reasons for delay or denial of surgery. *Obesity Surgery Journal*, 17, 1578–1583.