Measuring wellness among resident physicians

Samuel M. Keim, Mary Z. Mays, Jean M. Williams, Joyce Serido & Robin B. Harris

To cite this article: Samuel M. Keim, Mary Z. Mays, Jean M. Williams, Joyce Serido & Robin B. Harris (2006) Measuring wellness among resident physicians, Medical Teacher, 28:4, 370-374, DOI: 10.1080/01421590600625320

To link to this article: http://dx.doi.org/10.1080/01421590600625320

Published online: 03 Jul 2009.

Submit your article to this journal

Article views: 385

View related articles
Measuring wellness among resident physicians

SAMUEL M. KEIM, MARY Z. MAYS, JEAN M. WILLIAMS, JOYCE SERIDO & ROBIN B. HARRIS

University of Arizona Health Sciences Center, Tucson, USA

ABSTRACT Requirements to include professionalism in residency curricula have generated a substantial body of literature concerning the environments that fail to nurture professionalism. Local and national surveys provide evidence that a high prevalence of depersonalization and emotional exhaustion exists among residents and that clinical practice is impaired as a result of these factors. A group of 34 residents from ten residency programmes participated in the psychometric testing of a resident wellness assessment instrument that can be rapidly administered, scored, and interpreted. The Brief Resident Wellness Profile is composed of a Mood faces graphical rating item and a six-question subscale. The six-item subscale had good reliability (alpha = 0.83; r = 0.84), convergent validity (r = 0.63), discriminant validity (r = -0.37), and concurrent validity (p = 0.007). The Mood faces item had good convergent validity (r = 0.66), discriminant validity (r = -0.71), and concurrent validity (p = 0.008). The Brief Resident Wellness Profile appears to be a reliable and valid instrument that measures residents’ sense of professional accomplishment and mood and can be rapidly administered, scored, and interpreted.

Introduction

Requirements to include professionalism in residency curricula have generated a substantial body of literature. Comprehensive definitions of professionalism have been proposed (Lynch et al., 2004; van de Camp et al., 2004; Angoff, 2005). New curricula for teaching it have been developed (Haidet et al., 2005; Steinert et al., 2005). The types of methods used to assess it have been compared (Ginsburg et al., 2004; Smith et al., 2004; Stern et al., 2005; Veloski et al., 2005). Particular attention has been given to describing the environments that nurture it (Houy et al., 2000; Defoe et al., 2001; Vessey et al., 2002; Nuthalapaty et al., 2003; Gelfand et al., 2004; Mareiniss, 2004; Papp et al., 2004; Rosen et al., 2004) and the consequences of failing to do so (Shapiro & Lie, 2004; Thomas, 2004; Bellini & Shea, 2005). Local and national surveys have been conducted to estimate the prevalence of factors that are believed to facilitate or hinder the development of professionalism (Daugherty et al., 1998; Clever, 2002; Cohen, 2002; Collier et al., 2002; Shanafelt et al., 2002; Baldwin et al., 2004; Sargent et al., 2004). This literature provides ample evidence that a high prevalence of depersonalization and emotional exhaustion exists among residents and that clinical practice is impaired as a result of these factors.

Assessing resident wellness at the system level (rather than solely attempting to identify individual residents with current or incipient problems) should enable residency directors to create and sustain residency curricula and environments that nurture professionalism. Because of the intensity of graduate medical education, a significant obstacle to routinely and repetitively measuring resident wellness is the task and time burden of the testing itself. Other measurement issues surrounding the assessment of stress made it essential that the instrument has good psychometric properties (Hahn & Smith, 1999). An analysis of the literature in 2001 did not identify an instrument suitable for routine assessment of resident wellness, so we began developing a brief measure of resident wellness. A recent review of instruments used to measure constructs comprising professionalism also failed to identify a suitable instrument (Veloski et al., 2005). This report describes the development and preliminary validation of a resident wellness assessment instrument that can be rapidly administered, scored, and interpreted.

Method

In 2003, we began a longitudinal assessment of residents using a new instrument specifically developed to assess resident wellness. Standardized measures of job stress and mood were administered simultaneously in order to evaluate the validity of the new instrument. All assessments were self-administered in a single study questionnaire that also included demographic questions. Participants were emailed the questionnaire so that they could complete it at a time and place that was convenient and comfortable.

Participants

The University of Arizona Institutional Review Board approved the participation of human subjects in this research project. Volunteers were treated in accordance with the

Practice points

- Local and national surveys provide ample evidence that a high prevalence of depersonalization and emotional exhaustion exists among residents and that clinical practice is impaired as a result of these factors.
- The Brief Resident Wellness Profile is a psychometrically sound instrument that measures residents’ sense of professional accomplishment and mood and can be rapidly administered, scored, and interpreted.

Correspondence: Samuel M. Keim, PO Box 245057, University of Arizona Health Sciences Center, Tucson, AZ 85724-5057, USA, Email: sam@aemrc.arizona.edu
American Psychological Association’s (APA) ‘Ethical Principles of Psychologists and Code of Conduct’ (APA, 1992). Twenty residents volunteered to assist in creating a pool of potential items for the instrument. They represented six different residency programmes and postgraduate years one to five. Once the instrument was created, a random sample of 50 residents (from the 357 residents at the institution) were invited to participate in a one-year test of the instrument; 34 (68%) volunteered and completed the baseline questionnaire in spring 2003.

This group of 34 residents represented ten residency programmes: Anaesthesiology (2), Emergency Medicine (7), Family Practice (3), Medicine (5), Obstetrics/Gynaecology (5), Pathology (1), Paediatrics (5), Psychiatry (2), Radiology (2), and Urology (1). Six of the participants (18%) were completing their first year of residency, 11 (33%) participants their second year, and 16 (49%) participants their third year. Four individuals (12%) were graduates of the University of Arizona medical school; the remaining were graduates of schools in other states. Women made up 39% of the sample. When asked their racial identity, 27 individuals (82%) selected Caucasian, three (9%) selected Asian/Pacific Islander, one (3%) selected Native American, and two (6%) selected Other. One individual (3%) self-identified as being Hispanic. Ages ranged from 27 to 54 years (\(M = 31.6, S = 5.3\)). Ten individuals (30%) were single without children, one (3%) was a single parent, 11 (33%) were living with a spouse or partner, and 11 (33%) were living with a spouse or partner and children. Ten individuals (30%) reported working less than 50 h during the previous week, five (15%) 50 to 59 h, four (12%) 60 to 69 h, seven (21%) 70 to 80 h, and seven (21%) 80 h or more.

**Instruments**

_Brief Resident Wellness Profile (BRWP)._ The face validity of the new instrument was insured by engaging residents in the development of its profile. Items for the BRWP were derived from one-hour lasting structured interviews with 20 residents. Residents volunteered to be interviewed regarding their mood states and the behaviours they associated with their worst, normal, and peak performance during residency. Six items were developed from the issues most frequently raised as critically important ones (Table 1). These six items were scored on a scale of 1 (never) to 5 (always) and a total score was calculated by summing the ratings over the six items.

During the structured interview, residents were also asked to draw seven serial faces that reflected the mood states they had experienced and that included at least a mouth and eyes. Three descriptive anchors were provided ‘most happy you have felt’, ‘neutral—neither happy nor sad’, and ‘most sadness you have felt’. A professional biomedical graphic design artist used these hand drawings to create a Mood faces item that best represented the faces and points along the scale (Table 1). Responses to the mood item were scored on a scale of 0 (most unhappy face) to 6 (happiest face).

**Risk factors.** The contribution of personal stressors and long work hours to depersonalization and emotional exhaustion among residents has been well documented (Gelfand et al., 2004; Mareiniss, 2004; Papp et al., 2004; Thomas, 2004). In structured interviews, residents frequently cited stress related to personal relationships outside of work as a major contributor to their overall stress, their ability to cope with work-related stress, and their ability to maintain peak performance in the hospital. Thus, a simple dichotomous question about significant outside-of-work personal stress was included in the study questionnaire. The residents also frequently cited the number of hours worked as an important contributor to work-related stress. ‘Too much’ was described by all interviewees as greater than eighty hours per week, while ‘easy to handle’ was described by most as less than 50 hours per week. Two questions were derived from Table 1.

**Table 1. Brief Resident Wellness Profile.**

Below is a list of common feelings and attitudes that people experience. For each, fill in the circle corresponding to how often you have experienced or felt them _during the past week including today._

<table>
<thead>
<tr>
<th>Feeling/attitude</th>
<th>Rating scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never</td>
</tr>
<tr>
<td>1. Enthusiasm about career goals.</td>
<td>○</td>
</tr>
<tr>
<td>2. Willingness to do ‘whatever it takes’.</td>
<td>○</td>
</tr>
<tr>
<td>5. Feeling satisfied with progress so far in the residency.</td>
<td>○</td>
</tr>
<tr>
<td>6. Pleased with life overall.</td>
<td>○</td>
</tr>
</tbody>
</table>

Below is a series of faces that represent common mood states that residents experience. Fill in the circle corresponding to the face that best illustrates how you have been feeling _during the past week, including today._

![Faces](image-url)
these discussions: (1) ‘How many hours have you worked this past week?’ and (2) ‘How many hours per week have you been working on average?’ Residents chose from among five answers to the questions: <50, 50 to 59, 60 to 69, 70 to 79, and ≥80.

**Stress Profile (SP).** The SP (Western Psychological Services, Los Angeles, CA) was included to examine the validity of the BRWP as a measure of stress, coping, and well-being. This standardized instrument is used to identify behaviours that protect against or contribute to stress-related illnesses. The profile contains 123 items arranged in 14 subscales. Norms were developed on employees from manufacturing, aerospace, communications, and health care organizations. Two-week test-retest reliability and Cronbach coefficients exceeded 0.65. Scores on the profile successfully predict absenteeism, burnout, immune response, and job performance (Nowack, 1986, 1989, 1990, 1991). Four subscales of the 123-item profile were used: Stress, Cognitive hardiness, psychological well-being, and Response distortion bias. The six items of the stress subscale were scored on a scale of 1 (not at all) to 5 (extremely) and a total score was calculated by summing the ratings over the six items. The Cognitive hardiness subscale had 30 items; 17 items were scored positively on a scale of 1 (strongly agree) to 5 (strongly disagree) and 13 items were scored negatively on a scale of 1 (strongly agree) to 5 (strongly disagree). A total Cognitive hardiness score was calculated by summing the ratings over the 30 items. The 12 items of the Psychological well-being subscale were scored on a scale of 1 (never) to 5 (always) and a total score was calculated by summing the ratings over the 12 items. The Response distortion bias subscale had five items; four items were scored positively on a scale of 0 (false) to 1 (true) and one item was scored negatively on a scale of 0 (true) to 1 (false). A total Response distortion bias score was calculated by summing the ratings over the five items.

**Positive States of Mind (PSOM) questionnaire.** The PSOM questionnaire was included in order to examine the validity of the BRWP as a measure of mood and stress. This standardized six-item instrument is used to measure the ability to achieve and appreciate positive experiences, which are associated with perceived health and health promoting lifestyles. Norms were developed on undergraduates, athletes, pregnant women undergoing amniocentesis, and first year medical students. Cronbach coefficients exceeded 0.65. The four-week test-retest reliability coefficient was 0.46 consistent with its state orientation. Scores on the profile successfully predict moods, optimism, anxiety, somatic symptoms, and stress (Horowitz et al., 1988; Adler et al., 1998; Park & Adler, 2003). Items on the questionnaire were scored on a scale of 0 (unable to have it) to 3 (have it well) and a total score was calculated by summing the ratings over the six items.

**Statistical analysis**

Baseline data from the longitudinal project were used to assess the reliability and validity of the new instrument. All analyses were completed using SPSS software (SPSS, Inc. version 11.5.2.1, Chicago, IL). Simple frequencies and descriptive statistics were calculated to characterize the sample and scores on the individual instruments used in the study questionnaire. One individual scored 5 on the Response distortion bias subscale, indicating substantial bias. This individual was also the only individual who declined to answer the Mood item. These data were not used in any of the analyses. Results were considered to be statistically significant when \( p \leq 0.05 \).

Reliability coefficients were calculated using Cronbach alpha, Guttmann split-half, and equal-length Spearman-Brown methods. Construct validity was evaluated by factor analysis using a principal components method with varimax rotation (Kaiser normalization), analysing the correlation matrix, and extracting components with eigenvalues greater than 1.00. Convergent validity was assessed by calculating Pearson product-moment correlations between subscales designed to be highly and positively correlated. Discriminant validity was assessed by calculating Pearson product-moment correlations between subscales designed to be highly, but negatively correlated. Concurrent validity was assessed by comparing subgroups using the independent samples \( t \)-tests. Test bias was assessed using Pearson product-moment correlations, independent samples \( t \)-tests when two subgroups were compared, and one-way analyses of variance when more than two subgroups were compared.

**Results**

As shown in Table 2, scores on the instruments reflected the diversity of the sample mean scores, that were similar to those seen in other professional groups (Nowack, 1986, 1989, 1990, 1991; Horowitz et al., 1988; Adler et al., 1998). The data confirmed the prevalence results of other investigations (Collier et al., 2002; Baldwin & Daugherty, 2004; Gelfand et al., 2004; Sargent et al., 2004; Shanafelt et al., 2002; Thomas, 2004); 45% of residents reported high levels of stress (score > 15 on the stress subscale) and 39% reported negative mood on the Mood faces subscale (score < 4).

Despite the subscale’s brief nature, the six text items of the BRWP had good reliability, with a Cronbach alpha coefficient of 0.83, a Guttmann split-half coefficient of 0.84, and an equal-length Spearman-Brown coefficient of 0.84. Factor analysis identified two distinct factors that together accounted for 72% of the variance (Table 3). The first factor, which we termed ‘professional accomplishment’, included items 1, 2, 4, and 5; the second dimension, termed ‘mood’, included the two remaining items. The subscale had good convergent validity, as well, with correlations of 0.43 with Cognitive hardiness, 0.63 with Psychological well being, and 0.36 with Mood faces.

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress</td>
<td>15.1</td>
<td>3.0</td>
</tr>
<tr>
<td>Cognitive hardiness</td>
<td>106.0</td>
<td>15.6</td>
</tr>
<tr>
<td>Psychological well being</td>
<td>41.5</td>
<td>8.4</td>
</tr>
<tr>
<td>Response distortion bias</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td>Positive states of mind</td>
<td>11.7</td>
<td>3.6</td>
</tr>
<tr>
<td>Brief resident wellness profile</td>
<td>20.2</td>
<td>3.8</td>
</tr>
<tr>
<td>Mood faces</td>
<td>3.6</td>
<td>1.4</td>
</tr>
</tbody>
</table>

As shown in Table 2, scores on the instruments reflected the diversity of the sample mean scores, that were similar to those seen in other professional groups (Nowack, 1986, 1989, 1990, 1991; Horowitz et al., 1988; Adler et al., 1998). The data confirmed the prevalence results of other investigations (Collier et al., 2002; Baldwin & Daugherty, 2004; Gelfand et al., 2004; Sargent et al., 2004; Shanafelt et al., 2002; Thomas, 2004); 45% of residents reported high levels of stress (score > 15 on the stress subscale) and 39% reported negative mood on the Mood faces subscale (score < 4).
0.56 with Positive states of mind (p<0.01 in all cases). Discriminant validity was good with correlation of −0.37 with Stress (p=0.03). Concurrent validity was good as demonstrated by the relationship between the BRWP, workload, and stressors outside of work. Residents who reported working 60 h or less in the past week (n=15) had significantly higher BRWP scores, M=22.07, S=3.13 than those who had worked over 60 h (n=18), M=18.61, S=3.60 (t (31)=2.91, p=0.007). Residents who reported no significant stress in their personal lives outside of work (n=16) had significantly higher BRWP scores, M=22.38, S=2.83 than those who reported having significant stress (n=17), M=18.12, S=3.41 (t (31)=3.89, p<0.001). There was no evidence of test bias; i.e., BRWP scores were not influenced by simple demographic differences. Age was not significantly correlated with BRWP scores (r=−0.24, p>0.05). There were no significant differences in BRWP scores between men and women, between minorities and non-minorities, across different relationship categories, or between first, second, and third year residents (p>0.05 in all cases).

The single item Mood faces subscale also showed good convergent validity with correlations of 0.70 with Psychological well being and 0.66 with Positive states of mind (p<0.001 in both cases). Discriminant validity was also excellent with a correlation of −0.71 with Stress (p<0.001). Concurrent validity was good as demonstrated by the relationship between the Mood, workload, and stressors outside of work. Residents who reported working 60 h or less in the past week (n=15) had significantly higher Mood scores, M=4.27, S=0.80 than those who had worked over 60 h (n=18), M=3.00, S=1.57 (t (31)=2.83, p=0.008). Residents who reported no significant stress in their personal lives outside of work (n=16) had significantly higher Mood scores, M=4.13, S=0.89 than those who reported having significant stress (n=17), M=3.06, S=1.64 (t (31)=2.31, p=0.03). There was no evidence of test bias; i.e., the Mood faces subscale was not influenced by simple demographic differences. Age was not significantly correlated with Mood scores (r=−0.16, p>0.05). There were no significant differences in Mood scores between men and women, between minorities and non-minorities, across different relationship categories, or between first, second, and third year residents (p>0.05 in all cases).

**Discussion**

This report describes the development of the BRWP, a psychometrically sound instrument that measures residents’ sense of professional accomplishment and mood. Professional accomplishment was measured with four questions and mood with two questions and one graphic rating scale. The brevity of the instrument and its straightforward presentation allow for rapid administration, scoring, and interpretation. Analysis of baseline data from an ongoing longitudinal study of resident wellness demonstrated that the BRWP had good reliability and face, construct, convergent, discriminant, and concurrent validity. Future investigations will examine the utility of administering the BRWP via a web interface, the predictive validity of the instrument, patterns of change in wellness over the course of the residency programme, and the feasibility of using BRWP data to evaluate the effectiveness of changes in residency curricula and environments.

**Acknowledgements**

Portions of this work were supported by the Arizona Clinical Research Training Program, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, MD, grant #K30HL004519 and the University of Arizona Faculty Small Grants Program, grant # 210295.

**Notes on contributors**

SAMUEL M. KEIM, MD, is an Associate Professor of Emergency Medicine and the Director of the Emergency Medicine Residency program. His research interests are in medical education, evidence based medicine, wellness, and heat-related illnesses.

MARY Z. MAY, PhD, is a Research Associate Professor in the Department of Family and Community Medicine. She has broad experience in basic and applied behavioural and biomedical research both as an investigator and as a biostatistician.

JEAN M. WILLIAMS, PhD, is Professor of Psychology in the Department of Psychology. Her research interest has focused on the areas of psychology of injury, psychology of excellence, group dynamics and productivity, and exercise and mental health.

JOYCE SERIDO, PhD is a post-doctoral fellow at Cornell University, Ithaca, New York. Her research interests have focused on the effects of stress and decisions on careers, family relationships and personal well-being.

ROBIN B. HARRIS, PhD MPH is Director of the Division of Epidemiology in the Mel and Enid Zimmerman College of Public Health.

---

**Table 3. Factor loadings for Brief Resident Wellness Profile (n = 33).**

<table>
<thead>
<tr>
<th>Feeling/attitude during the past week, including today?</th>
<th>Professional accomplishment</th>
<th>Mood</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Feeling satisfied with progress so far in the residency.</td>
<td>0.85</td>
<td>0.10</td>
</tr>
<tr>
<td>1. Enthusiasm about career goals.</td>
<td>0.82</td>
<td>0.29</td>
</tr>
<tr>
<td>2. Willingness to do ‘whatever it takes’.</td>
<td>0.76</td>
<td>0.15</td>
</tr>
<tr>
<td>4. Feeling confident with challenging patients.</td>
<td>0.71</td>
<td>0.20</td>
</tr>
<tr>
<td>3. Feeling calm, poised, and stress-free.</td>
<td>0.07</td>
<td>0.95</td>
</tr>
<tr>
<td>6. Pleased with life overall.</td>
<td>0.50</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Bold highlights the factor loadings relevant to the proposed dimensions in each column.
Her broad research interests have recently focused on cancer prevention, health care delivery and public health education.

References


