



Resident Physician Well-Being and Assessments of Their Knowledge and Clinical Performance

Thomas J. Beckman, MD¹, Darcy A. Reed, MD, MPH¹, Tait D. Shanafelt, MD¹,
and Colin P. West, MD, PhD^{1,2}

¹Department of Internal Medicine, College of Medicine, Mayo Clinic, Rochester, MN, USA; ²Department of Health Sciences Research, Division of Biomedical Statistics and Informatics, College of Medicine, Mayo Clinic, Rochester, MN, USA.

BACKGROUND: Medical knowledge and clinical performance ratings are major criteria for assessing the competence of resident physicians. However, these assessments may be influenced by residents' mental health. The relationship between residents' well-being and empathy and assessments of their global performance remains unclear.

OBJECTIVE: To determine whether resident well-being and empathy are associated with assessments of their medical knowledge and clinical performance.

DESIGN, SETTING, AND PARTICIPANTS: We studied 730 clinical performance assessments completed by peers, supervising residents, and allied health professionals; 193 mini-clinical evaluation exercise (mini-CEX) evaluations; and 260 in-training examinations (ITE) of Mayo Clinic internal medicine residents between January 2009 and August 2010. Multivariate generalized estimating equations were used to evaluate associations between residents' well-being and empathy and assessments of their knowledge and clinical performance.

MEASUREMENTS: Independent variables were empathy using the Interpersonal Reactivity Index (IRI), burnout using the Maslach Burnout Inventory, depression using a standardized two-question screening instrument, and quality of life using a Linear Analog Self-Assessment item and the Medical Outcomes Study Short Form (SF-8). Dependent variables were mini-CEX, ITE, and the validated six-item Mayo clinical performance assessment.

RESULTS: 202 residents (64.7%) provided both well-being and at least one category of assessment data. In multivariate models, residents' scores on the IRI empathy measure of "the tendency to adopt the psychological view of others" were associated with higher peer ratings on "desirability as a physician for a family member" (beta=0.023, 95% CI=0.007–0.039, p=0.004). Additionally, burnout was associated with higher supervisor ratings of communication (beta=0.309, 95% CI=0.100–0.517, p=0.004). There were no observed associations between ITE or mini-CEX scores and resident quality of life, burnout, fatigue, depression, or empathy.

Electronic supplementary material The online version of this article (doi:10.1007/s11606-011-1891-6) contains supplementary material, which is available to authorized users.

Received May 20, 2011

Revised August 29, 2011

Accepted September 9, 2011

Published online September 24, 2011

CONCLUSIONS: Most dimensions of resident well-being were not associated with residents' knowledge scores and assessments of their clinical performance by other members of the health care team, which supports the trustworthiness of these measures. Nonetheless, correlations of resident empathy and burnout with assessments completed by peers and supervising residents suggest that some ratings of residents may be influenced by interpersonal factors.

KEY WORDS: graduate medical education; well-being; empathy; assessment.

J Gen Intern Med 27(3):325–30

DOI: 10.1007/s11606-011-1891-6

© Society of General Internal Medicine 2011

BACKGROUND

Medical knowledge and clinical performance ratings are major criteria for assessing the competence of resident physicians,¹ so it is essential that these standards are trustworthy. Assessments of medical knowledge and clinical performance should primarily reflect residents' abilities to care for patients, but these assessments may be influenced by other factors. For example, research has shown that resident empathy affects assessments of faculty members,² female teachers receive lower ratings than males with similar skills,³ and assessment scores for a single instrument may vary substantially between learning environments within the same institution.⁴

It has been theorized that physician health and well-being could have an impact on medical knowledge acquisition.⁵ While some studies have linked resident well-being and medical knowledge,^{6,7} other research has failed to demonstrate this relationship.^{8–10} It is unknown whether resident well-being shapes overall assessments of residents' competency.

Research has shown that internal medicine residents' well-being is affected by features of the work environment such as duty-hours¹¹ and the perception of experiencing medical errors.¹² Because it is known that learning environments^{11,12} and interpersonal relationships^{13–16} play crucial roles in learning, we postulated that residents' well-being and empathy would influence assessments of their medical knowledge and clinical performance.

Studies have indicated that well-being affects residents' attitudes towards patients and the quality of care that they provide,^{11,12,17} but no studies have examined interactions

between residents' well-being and the global assessments they receive from other members of the health care team. Therefore, we used a prospective longitudinal study design to investigate the hypothesis that resident well-being (measured by the Maslach Burnout Inventory, Linear Analog Self-Assessment item, Medical Outcomes Study Short Form health survey, and a standardized two-question depression screen) and empathy (measured by the Interpersonal Reactivity Index) are associated with numerous dimensions of competency including assessments of knowledge on the medical in-training examination (ITE), clinical performance on the mini-clinical evaluation exercise (mini-CEX), and multi-source assessments by peers, supervising residents, and allied health professionals.

METHODS

Learning Environment and Participants. This study involved Mayo internal medicine residents in training between January 2009 and August 2010, and included scores from all knowledge and clinical performance assessments performed on the 264 resident physicians enrolled during this time period in the Mayo Clinic Internal Medicine Well-being (IMWELL) Study, which is described below. This study was approved by the Mayo Institutional Review Board.

The Mayo IMWELL Study. Resident characteristics were obtained from the Mayo Clinic-Rochester IMWELL study, a longitudinal study of resident physician well-being. Since the 2003 academic year, all categorical and preliminary residents in the Mayo Clinic Rochester Internal Medicine Residency program have been invited to participate in the IMWELL study during their first-year orientation. For the time period of this study, 264 of 312 (84.6%) eligible residents volunteered to participate. All residents provided written consent and were surveyed at regular intervals throughout their residency training. An instrument that measures quality of life (QOL) was administered quarterly, and instruments that measure burnout, empathy, and depression were administered biannually. To maintain anonymity, study participants' identities were blinded by using numerical codes during data collection and analyses.

Instruments Comprising the Mayo IMWELL Study. The IMWELL study utilized a linear analog self-assessment (LASA) scale of QOL and survey items from the Maslach Burnout Inventory (MBI), Interpersonal Reactivity Index (IRI), Medical Outcomes Study Short Form (SF-8) health survey, and a depression screen by Spitzer et al. Notably, these instruments are supported by sources of validity evidence^{18,19} that are essential to medical education studies.^{20,21}

The LASA for measuring QOL is a single item with scores ranging from 0 (as bad as it can be) to 10 (as good as it can be). LASA QOL scores have been validated in varied populations including the general public,²² cancer patients,^{23,24} and physicians.²⁵

The MBI is a 22-item instrument with Likert scales ranging from 0 (never) to 6 (daily).²⁶ Content validity has been demonstrated by reviewing established scales and surveying professionals who are at risk of experiencing burnout, including physicians.²⁶ Many studies have shown that the MBI is an effective measure of burnout in resident physicians.^{12,27-30} Factor analysis demonstrated that the MBI consists of three dimensions: depersonalization, emotional exhaustion, and sense of low personal accomplishment.²⁶ As has been done in previous studies of physicians, we considered residents with a high score on either the depersonalization or emotional exhaustion subscale as having at least one manifestation of professional burnout.^{25,27} Additionally, the MBI has high internal consistency, acceptable test-retest reliability, moderate correlation with other measures of burnout, and poor correlation with constructs that are likely confounded with burnout.²⁶

The IRI is a 28-item instrument for measuring empathy with Likert scales ranging from 0 (does not describe me well) to 4 (describes me well). Factor analysis revealed four dimensions: perspective-taking, personal distress, empathic concern, and fantasy.³¹ These seven-item subscales correspond to the two-dimensional cognitive and emotive model, and may be evaluated separately.^{31,32} A model focusing on empathic concern and perspective-taking has proven especially useful when evaluating empathy among resident physicians.^{17,30,33,34} On this basis, these two subscales were assessed in this study. Additional validity evidence for the IRI includes good internal consistency,^{32,35,36} and significant correlations between IRI subscales and other recognized measures of empathy.^{36,37}

The Medical Outcomes Study Short Form (SF-8) health survey has eight items with 5- and 6-point Likert scales.³⁸ Content validity is supported in that survey items represent ideas that are commonly included in widely used health measures.³⁸ The SF-8 generates scores that are assigned to domains of mental and physical health.³⁸ Further SF-8 validity evidence includes score reliability, high correlation with existing measures of the same concepts,³⁸ and convergence between measurements of patients with migraine and those with other conditions.³⁹

Spitzer et al. described a depression screening method consisting of two questions: "During the past month, have you often been bothered by feeling down, depressed and hopeless?" and "During the past month, have you often been bothered by little interest or pleasure in doing things?"⁴⁰ These two questions perform well in screening for depression relative to several widely used depression inventories, including the Beck Depression Inventory, the Center for Epidemiological Studies Depression Scale, and the Medical Outcomes Study depression measure.⁴¹ Spitzer et al.'s screening questions have been used to identify depression in various populations, including resident physicians.^{11,27}

Outcome Measures. Strong validity evidence supports the outcome measures used in this study. ITE scores, the validity and use of which are well described,^{42,43} have been shown to correlate strongly with resident conference attendance and self-directed reading,^{44,45} and to have no association with resident physician empathy.¹⁰ The mini-CEX has impressive validity and reliability as demonstrated by previous studies at the Mayo Clinic^{46,47} and elsewhere.⁴⁸⁻⁵² Our institution uses

the traditional version of the mini-CEX, except that the items are on 5-point scales.

Clinical performance assessments of resident physicians at the Mayo Clinic are comprised of forms completed by peers, senior medical residents, and non-physician professionals. After critically analyzing the items for content most plausibly related to resident well-being, the following items (scale: 1=needs improvement, 3=average, 5=top 10%) were chosen for this study: (1) desirability as a physician for one of your family members, (2) desirability as a future co-worker or team member, (3) effectiveness and completeness of sign-outs, (4) coverage of cross-cover issues and completeness of tasks when on call, (5) demonstrates empathy and compassion for patients, and (6) communication skills with patients, family, allied health, and other providers. Content validity for these clinical performance assessments is based on assessment elements that are represented in previously published instruments and were selected by experts with experience in scale design. A factor analytic study revealed that several items within these Mayo clinical performance assessments are multi-dimensional and have excellent internal consistency reliability.⁵³

Data Analysis. A repeated measures design, analyzed using multivariate generalized estimating equations, was employed to evaluate associations between resident clinical performance assessments, mini-CEX evaluations and ITE examinations, and residents' QOL, burnout, empathy, and depression over the 4 points in time. Scores from the clinical performance assessment items were evaluated individually and also averaged within assessor group to form an overall score ranging from 1 to 5. Covariates included resident well-being (QOL, burnout, depression), empathy, gender, year of training, program (categorical or preliminary), debt, relationship status (single, married, divorced, partner), and children (yes or no). Univariate associations were examined, and a multivariate model was developed using standard forward and backward stepwise selection techniques. The threshold for statistical significance was set at $p < 0.01$ to account for multiple comparisons. The study sample of 202 residents provided 80% power for a medium-to-small Cohen's f^2 effect size of 0.04 for a univariate association between clinical performance assessment scores, mini-CEX evaluations, and ITE examinations, and any well-being or empathy variable. Statistical analyses were conducted using SAS version 9.1 (SAS Institute Inc., Cary, NC).

RESULTS

We studied 730 clinical performance assessments completed by peers, supervisors, and allied health professionals for Mayo Clinic internal medicine residents in January 2009, August 2009, January 2010, and August 2010. We also examined 193 mini-clinical evaluation exercise (mini-CEX) evaluations and 260 in-training examinations (ITE) during the same time frame. Of 312 eligible residents, 202 (64.7% of all eligible) provided well-being and at least one category of assessment data for this study. Demographic characteristics for this sample are shown in Table 1. Data for responders were similar

Table 1. Baseline Characteristics of Resident Physicians Providing Both Well-Being and Evaluation Data from January 2009 Through August 2010

Variable	Level	Eligible	n	% of sample
Year of training	PGY 1	216	147	72.8
	PGY 2	48	28	13.9
	PGY 3	48	27	13.4
Age	24–30 years	261	147	85.5
	≥31 years	51	25	14.5
Sex	Male	188	116	57.4
	Female	124	86	42.6
Program	Categorical	240	155	76.7
	Preliminary	72	47	23.3

Notes: Data reflect status at first eligible time point

to those of non-responders on measured factors including year of training, age, sex, and program type (categorical or preliminary). In addition, United States Medical Licensing Examination Step 1 and Step 2 Clinical Knowledge (CK) scores were similar for responders and non-responders (Step 1 mean scores 232.3 vs 231.2, respectively, $p = 0.74$; Step 2 CK mean scores 241.3 vs 236.1, respectively, $p = 0.10$). As the overall results did not differ for categorical and preliminary residents, data were pooled across these categories. Baseline well-being, empathy, assessment, mini-CEX, and ITE scores at the start of the current study are shown in Table 2. In addition, response rates for each outcome are detailed in Table 2.

Univariate associations of measures of well-being and empathy with summary scores in each resident performance assessment domain are shown in an on-line appendix. In multivariate models, there were no statistically significant associations between resident ITE or mini-CEX scores and QOL, burnout, depression, empathy, or demographic characteristics. Both forward and backward model selection approaches yielded the same results. However, residents' scores on the IRI Perspective Taking scale, a measure of "the tendency to adopt the psychological view of others," were associated with higher peer ratings on "desirability as a physician for a family member" (multivariate $\beta = 0.023$, 95% CI = 0.007–0.039, $p = 0.004$). Consequently, a 5-point increase in this empathy score was associated with a small but statistically significant 0.12-point increase in residents' ratings by their peers as desirable physicians, as shown in Table 3. Additionally, having at least one manifestation of professional burnout was associated with higher resident supervisor ratings of communication with patients, families, allied health, and other providers (multivariate $\beta = 0.309$, 95% CI = 0.100–0.517, $p = 0.004$). Hence, burnout was associated with a 0.3-point increase in resident communication score, as shown in Table 3.

DISCUSSION

To our awareness, this is the first study to show that multiple dimensions of resident performance assessment are not significantly influenced by various aspects of well-being including QOL and depression. Nonetheless, the detected correlations between resident empathy and burnout and

Table 2. Initial Well-Being, Empathy, Assessment, Mini-CEX, and ITE Scores for Resident Physicians Providing Data from January 2009 to August 2010

Variable	Metric (scale)	n (Eligible)*	n (Responses)	Mean (SD)	Range
Quality-of-life	LASA overall QOL (0–10)	312	202	6.73 (1.77)	2–10
Health survey	SF-8 (Mental) (0–100)	312	196	46.57 (9.10)	12.2–60.4
	SF-8 (Physical) (0–100)	312	196	53.15 (5.99)	33.7–66.4
Burnout	MBI-DP (0–30)	312	202	8.81 (6.34)	0–28
	MBI-EE (0–54)	312	202	23.01 (11.30)	1–52
	MBI-PA (0–48)	312	202	37.86 (6.46)	17–48
	Positive 2-item screen, %	312	202	35.15 (47.74)	N/A
Depression	Positive 2-item screen, %	312	202	35.15 (47.74)	N/A
Empathy	IRI-PT (0–28)	312	198	19.44 (4.24)	7–28
	IRI-EC (0–28)	312	198	20.71 (4.51)	4–28
Baseline global resident ratings	Peer average (1–5)	312	185	4.13 (0.48)	2.1–5
	Supervisor average (1–5)	216	128	3.99 (0.64)	1.5–5
	Allied health average (1–5)	312	155	4.00 (0.52)	3–5
Clinical skill	Mini-CEX (1–5)	240	131	3.78 (0.12)	2.54–5
Knowledge	ITE overall score (0–100)	240	160	64.88 (8.00)	47–86

Note: variable completion rates for each category of assessment data are due to incomplete surveys, missing evaluation measures at the same point of time, and the fact that not all rotations have clinical evaluations

*312 residents were in training during the time period of this study. This total included two classes of 48 senior residents each who therefore did not have supervisory resident evaluations, resulting in 216 residents eligible for these evaluations. In addition, the three classes of 24 preliminary residents each did not have outpatient clinic and therefore did not have mini-CEX evaluations. These residents also do not take the ITE, resulting in 240 residents eligible for these evaluations

assessments completed by other doctors suggest that observation-based ratings of residents may be influenced by both intrapersonal and interpersonal factors.

Resident well-being has received national attention because the Accreditation Council for Graduate Medical Education has recently mandated further restrictions on resident duty hours,⁵⁴ based on the assumption that these restrictions will enhance patient care by improving resident fatigue and well-being. Consequently, it is important to examine the potential relationships between resident well-being and assessments of resident competency. Furthermore, many residency programs assess resident performance on tests of knowledge and clinical skill using the ITE and mini-CEX, respectively. Thus, it is encouraging that we identified no associations between resident well-being and performance on the ITE and mini-CEX. Our findings build upon research that has demonstrated validity of these measures,^{43,46–52} while underscoring the fact

that these measures should not be viewed as reliable means for detecting variations in resident empathy and well-being. Our findings also suggest that increasing resident well-being through duty hour restrictions may not be a comprehensive, stand-alone strategy for improving resident performance and enhancing patient care.

It has been theorized that resident performance is partly determined by depression, burnout, and distress related to the work environment.⁵ Girard et al. found that American Board of Internal Medicine (ABIM) Certification Examination score variation among internal medicine residents was largely attributable to residents' psychological states.⁶ Filho et al. have shown that anesthesiology residents' knowledge of basic science was associated with academic performance anxiety, but was not associated with QOL.⁹ However, other research has identified no relationship between resident well-being and knowledge on specific topics or standardized examinations.⁸

Table 3. Univariate Associations Between Well-Being and Empathy, with Desirability as a Physician and Effective Communication

Variable	Metric	Desirability as a physician			Effective communication		
		Beta [†]	95% CI	p-value	Beta [†]	95% CI	p-value
Quality-of-life	LASA overall QOL	0.0267	-0.02, 0.07	0.28	-0.0286	-0.07, 0.02	0.21
Health survey	SF-8 (Mental)	0.0006	-0.01, 0.01	0.88	-0.0016	-0.01, 0.01	0.80
	SF-8 (Physical)	0.0013	-0.01, 0.01	0.84	-0.0011	-0.02, 0.01	0.90
Burnout	Yes/no	0.0631	-0.20, 0.08	0.38	0.3046	0.10, 0.51	0.0039†
	MBI-DP	-0.0006	-0.01, 0.01	0.93	0.0239	0.0066, 0.0412	0.0068
	MBI-EE	-0.0028	-0.01, 0.004	0.43	0.0090	0.0004, 0.0177	0.0406
	MBI-PA	-0.0032	-0.01, 0.01	0.47	0.0020	-0.01, 0.02	0.81
Depression	2-item screen	0.0852	-0.06, 0.23	0.24	-0.2443	-0.4429, -0.0456	0.0160
Empathy	IRI-PT	0.0221	0.0061, 0.0380	0.0066†	-0.0047	-0.03, 0.02	0.73

Notes: [†]Parameter estimates reflect the change in either peer rating of desirability as a physician for a family member or supervising residents' rating of effective communication with patients, families, allied health, and other providers associated with a 1-unit increase in each metric. For example, a 1-point increase in IRI-PT was associated with a statistically significant 0.0221-point increase in peer rating of desirability as a physician for a family member, so that a 10-point increase in IRI-PT was associated with a statistically significant 0.221-point increase in peer rating of desirability as a physician for a family member

†These items remained statistically significant after multivariate adjustment. The level of statistical significance was set at $p=0.01$ to account for multiple comparisons

Likewise, the current study found no association between resident well-being and medical knowledge on the ITE. A potential explanation for this lack of association is that an insufficient number of residents in our study sample displayed the extremes of low well-being that would be required to negatively impact on the acquisition or display of medical knowledge; indeed, West et al. identified an association between resident well-being and medical knowledge in a much larger sample with wider ranges in well-being scores.⁷ Another explanation for our negative findings may be that the ITE—as opposed to the ABIM certification—is a lower stakes examination and intended only for formative feedback. This may be the reason why Girard et al. identified a relationship between resident mood disorders and score variation in the ABIM certification examination.⁶

We found an association between resident empathy and desirability as a physician. In particular, residents who scored high on the IRI Perspective-Taking scale, which is “the tendency to adopt the psychological views of others,” were considered to be desirable physicians for a family member. It is common wisdom that to be entrusted with the care of a colleague’s loved one is the highest compliment. This finding also reflects two core virtues of medicine, which include empathy and compassion.^{55,56} Additionally, this finding extends previous studies that have shown relationships between scores on the IRI Perspective-Taking scale and higher performance among medical students,⁵⁷ as well as decreased likelihood of medical errors by residents.¹²

We observed that supervising residents perceived interns with higher burnout to have *better* communication with patients, families, allied health, and other providers. Although this finding could seem counterintuitive, many physicians who experience burnout may sustain high levels of professional achievement for long durations. Furthermore, the most dedicated physicians might be more likely to place professional duties—including the time-consuming task of effectively communicating with patients, family members, support staff, and colleagues—above all other aspects of personal life.⁵⁸ Therefore, such physicians could be viewed favorably by supervisors in the workplace, even though the personal aspects of these physicians’ lives may suffer. Ultimately, this finding should prompt residency programs to reflect on the optimal balance between patient care responsibilities and resident burnout.

This study has limitations. It involved participation by only 64.7% of eligible residents. Of the 264 of 312 (84.6%) residents volunteering for the IMWELL study, there were additional missing data due to incomplete surveys, missing evaluation measures at the same point in time as well-being assessment, and the fact that not all rotations have clinical evaluations. Nonetheless, the participation rate in this study was favorable relative to that typically seen in physician studies.^{59,60} Our study sample was largely comprised of first-year residents and resident work-load—which may influence well-being—varies by PGY year; yet, our study results were adjusted for PGY level in the multivariate analyses, and none of the results were significantly different for the various PGY levels. Since this was a single-institution study, one should generalize the findings to other settings with caution. However, the range of well-being and empathy scores in this study is similar to those reported in previous studies at other institutions.^{28,33,34} Furthermore, the outcomes variables in this study for medical knowledge (ITE) and clinical skill (mini-CEX) are widely used among US

residency programs, which should broaden the importance and relevance of the study findings. While the results may have been affected by non-response bias, data for responders were similar to those of non-responders on measured demographic factors and USMLE scores. The small but statistically significant associations between resident empathy and burnout and assessments by other physicians might be viewed as clinically insignificant; however, the actual range of scores among residents at the Mayo Clinic for communications competency is very narrow, so the observed 0.3-point change in communication score with the presence of burnout (yes/no) could have a substantial impact on a Mayo resident’s relative standing within this competency. Additionally, the positive associations between resident empathy and burnout and assessments by other physicians persisted after multivariate adjustments that incorporated a number of potential confounding covariates, so we believe that these findings are educationally relevant and add new knowledge regarding the relationships between resident empathy and desirability as a physician, as well as between resident burnout and communication. Finally, we acknowledge that this study did not examine the influence of several potential confounders, including residents’ learning styles, personality types, and major life events.

In this study sample, multiple dimensions of resident performance were generally not influenced by various aspects of well-being, which lends credibility to standardized measures of knowledge (ITE) and clinical performance (mini-CEX). However, the lack of association between well-being and medical knowledge requires further study involving high-stakes examinations. The association between empathy and desirability as a doctor for a family member—which is widely accepted as a characteristic of excellence among physicians—suggests the need to emphasize the identification and promotion of empathy in medical learners. However, the positive association between burnout and the perception of excellent clinical performance should stimulate discussion about the best ways to engage residents in meaningful clinical experiences without compromising their overall well-being.

Acknowledgments: Presented at the Association for Medical Education in Europe Meeting in Vienna, Austria, on August 29, 2011. This study was supported by a Mayo Program in Professionalism and Ethics (PPE-5) Award.

Conflict of Interest: None disclosed.

Corresponding Author: Thomas J. Beckman, MD; Department of Internal Medicine, College of Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN, USA (e-mail: Beckman.Thomas@mayo.edu).

REFERENCES

1. ACGME. ACGME outcome project. 2004; www.acgme.org/outcome/.
2. Beckman TJ, Reed DA, Shanafelt TD, West CP. Impact of resident well-being and empathy on assessments of faculty physicians. *J Gen Intern Med.* 2010;25:52–6.
3. Beckman TJ, Mandrekar JN. The interpersonal, cognitive and efficiency domains of clinical teaching: construct validity of a multidimensional scale. *Med Educ.* 2005;39:1221–9.
4. Beckman TJ, Cook DA, Mandrekar JN. Factor instability of clinical teaching assessment scores among general internists and cardiologists. *Med Educ.* 2006;40:1209–16.

5. **Mitchell M, Srinivasan M, West DC, et al.** Factors affecting resident performance: development of a theoretical model and a focused literature review. *Acad Med.* 2005;80:376-9.
6. **Girard DE, Hickam DH.** Predictors of clinical performance among internal medicine residents. *J Gen Intern Med.* 1991;6:150-4.
7. **West C, Shanafelt TD, Kolars JC.** Quality of life, burnout, educational debt, and medical knowledge among internal medicine residents. *JAMA.* 2011;306:952-60.
8. **West CP, Shanafelt TD, Cook DA.** Lack of association between resident doctors' well-being and medical knowledge. *Med Educ.* 2010;44:1224-31.
9. **de Oliveira Filho GR, Vieira JE.** The relationship of learning environment, quality of life, and study strategies measures to anaesthesiology resident academic performance. *Anesth Analg.* 2007;104:1467-72.
10. **West CP, Huntington JL, Huschka MM, et al.** A prospective study of the relationship between medical knowledge and professionalism among internal medicine residents. *Acad Med.* 2007;82:587-92.
11. **Goitein L, Shanafelt TD, Wipf JE, Slatore CG, Back AL.** The effects of work-hour limitations on resident well-being, patient care, and education in an internal medicine residency program. *Arch Intern Med.* 2005;165:2601-6.
12. **West CP, Huschka MM, Novotny PJ, et al.** Association of perceived medical errors with resident distress and empathy. *JAMA.* 2006;296:1071-8.
13. **Beckman TJ, Lee MC.** Proposal for a collaborative approach to clinical teaching. *Mayo Clin Proc.* 2009;84:339-44.
14. **Williams GC, Wiener MW, Markakis KM, Reeve JM, Deci EL.** Medical students' motivation for internal medicine. *J Gen Intern Med.* 1994;9:327-33.
15. **Wortham S.** The interdependence of social identification and learning. *Am Educ Res J.* 2004;41:715-50.
16. **Haidet P, Stein HF.** The role of the student-teacher relationship in the formation of physicians. The hidden curriculum as process. *J Gen Intern Med.* 2006;21:S16-20.
17. **Shanafelt TD, West C, Zhao X, et al.** Relationship between increased personal well-being and enhanced empathy among internal medicine residents. *J Gen Intern Med.* 2005;20:559-64.
18. **Messick S.** Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *Am Psychol.* 1995;50:741-9.
19. **Downing SM.** Validity: on the meaningful interpretation of assessment data. *Medical Education.* 2003;37:830-7.
20. **Beckman TJ, Cook DA, Mandrekar JN.** What is the validity evidence for assessments of clinical teaching? *J Gen Intern Med.* 2005;20:1159-64.
21. **Beckman TJ, Ghosh AK, Cook DA, Erwin PJ, Mandrekar JN.** How reliable are assessments of clinical teaching? A review of the published instruments. *J Gen Intern Med.* 2004;19:971-7.
22. **Gudex C, Dolan P, Kind P, Williams A.** Health state valuations from the general public using the visual analog scale. *Qual Life Res.* 1996;5:521-31.
23. **Spitzer WO, Dobson AJ, Hall J, et al.** Measuring the quality of life in cancer patients: a concise QL-index for use by physicians. *J Chron Dis.* 1981;34:585-97.
24. **Rummans TA, Clark MM, Sloan JA, et al.** Impacting quality of life for patients with advanced cancer with a structured multidisciplinary intervention: a randomized controlled trial. *J Clin Oncol.* 2006;24:635-42.
25. **Shanafelt TD, Novotny P, Johnson ME, et al.** The well-being and personal wellness promotion strategies of medical oncologists in the North Central Cancer Treatment Group. *Oncology.* 2005;68:23-32.
26. **Maslach C, Jackson SE, Leiter MP.** *Maslach Burnout Inventory Manual* (3rd ed). Palo Alto, CA: Consulting Psychologists Press; 1996.
27. **Shanafelt TD, Bradley KA, Wipf JE, Back AL.** Burn-out and self-reported patient care in an internal medicine residency program. *Ann Intern Med.* 2002;136:358-67.
28. **Thomas NK.** Resident burnout. *JAMA.* 2004;292:2880-9.
29. **Gopal R, Glasheen JJ, Miyoshi TJ, Prochazka AV.** Burnout and internal medicine resident work-hour restrictions. *Arch Intern Med.* 2005;165:2595-600.
30. **Rosen IM, Gimotty PA, Shea JA, Bellini LM.** Evolution of sleep quality, sleep deprivation, mood disturbances, empathy, and burnout among interns. *Acad Med.* 2006;81:82-5.
31. **Davis M.** A multidimensional approach to individual differences in empathy. *JSAS Catalog Selected Documents Psychol.* 1980;10:85.
32. **Davis M.** Measuring individual differences in empathy: evidence for a multi-dimensional approach. *J Pers Soc Psychol.* 1983;44:113-26.
33. **Bellini LM, Baime M, Shea JA.** Variation of mood and empathy during internship. *JAMA.* 2002;287:3143-6.
34. **Bellini LM, Shea JA.** Mood change and empathy decline persist during three years of internal medicine training. *Acad Med.* 2005;80:164-7.
35. **Yarnold PR, Bryant FB, Nightingale SD, Martin GJ.** Assessing physician empathy using the Interpersonal Reactivity Index: A measurement model and cross-sectional analysis. *Psychology, Health, and Medicine.* 1996;1:207-21.
36. **Bryant FB, Yarnold PR, Grimm LG.** Toward a measurement model of the Affect Intensity Measure: a three-factor structure. *Journal of Research in Personality.* 1996;30:223-47.
37. **Lawrence EJ, Shaw P, Baker D, Baron-Cohen S, David AS.** Measuring empathy: reliability and validity of the empathy quotient. *Psychological Medicine.* 2004;34:911-24.
38. **Ware JE, Kosinski M, Dewey JE, Gandek B.** How to score and interpret single-item health status measures: a manual for users of the SF-8 Health Survey. Lincoln, RI: QualityMetric Incorporated; 2001.
39. **Turner-Bowker D, Bayliss M, Ware JEJ, Kosinski M.** Usefulness of the SF-8 health survey for comparing the impact of migraine and other conditions. *Qual Life Res.* 2003;12:1003-12.
40. **Spitzer RL, Williams JB, Kroenke K, et al.** Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD study. *JAMA.* 1994;272:1749-56.
41. **Whooley MA, Avins AL, Miranda J, Browner WS.** Case-finding instruments for depression. Two questions are as good as many. *J Gen Intern Med.* 1997;12:439-45.
42. **Garibaldi RA, Trontell MC, Waxman H, et al.** The in-training examination in internal medicine. *Ann Intern Med.* 1994;121:117-23.
43. **Garibaldi RA, Subhiyah R, Moore ME, Waxman H.** The in-training examination in internal medicine: an analysis of resident performance over time. *Ann Intern Med.* 2002;137:505-10.
44. **McDonald FS, Zeger SL, Kolars JC.** Factors associated with medical knowledge acquisition during internal medicine residency. *J Gen Intern Med.* 2007;22:962-8.
45. **McDonald FS, Zeger SL, Kolars JC.** Associations of conference attendance with internal medicine in-training examination scores. *Mayo Clin Proc.* 2008;83:449-53.
46. **Cook DA, Beckman TJ.** Does scale length matter? A comparison of nine- versus five-point rating scales for the mini-CEX. *Adv Health Sci Educ Theory Pract.* 2009;14:655-64.
47. **Cook DA, Beckman TJ, Mandrekar JN, Pankratz VS.** Internal structure of mini-CEX scores for internal medicine residents: factor analysis and generalizability. *Adv Health Sci Educ.* 2010;15:633-45.
48. **Norcini JJ, Blank LL, Duffy FD, Fortna GS.** The mini-CEX: a method for assessing clinical skills. *Ann Intern Med.* 2003;138:476-81.
49. **Kogan JR, Bellini LM, Shea JA.** Feasibility, reliability, and validity of the mini-clinical evaluation exercise (mCEX) in a medicine core clerkship. *Acad Med.* 2003;78(10 suppl):S33-5.
50. **Hatala R, Ainslie M, Kassen BO, Mackie I, Roberts JM.** Assessing the mini-clinical evaluation exercise in comparison to a national specialty examination. *Med Educ.* 2006;40:950-6.
51. **Holmboe ES, Hout S, Chung J, Norcini JJ, Hawkins RE.** Construct validity of the mini-clinical evaluation exercise (miniCEX). *Acad Med.* 2003;78:826-30.
52. **Durning SJ, Cation LJ, Markert RJ, Pangaro L.** Assessing the reliability and validity of the mini-clinical evaluation exercise for internal medicine residency training. *Acad Med.* 2002;77:900-4.
53. **Beckman TJ, Mandrekar JN, Engstler GJ, Ficalora RD.** Determining the reliability of clinical assessment scores in real time. *Teach Learn Med.* 2009;21:188-94.
54. **ACGME.** Common Program Requirements 2011. http://www.acgme.org/acwebsite/home/Common_Program_Requirements_07012011.pdf. Accessed August 22, 2011.
55. **Pellegrino ED.** Toward a reconstruction of medical morality. *The American Journal of Bioethics.* 2006;6:65-71.
56. **Rosenow EC III.** Recertifying in the art of medicine: what I would tell young physicians. *Mayo Clin Proc.* 2000;75:865-8.
57. **Stratton TD, Elam CL, Murphy-Spencer AE, Quinlivan SL.** Emotional intelligence and clinical skills: preliminary results from a comprehensive clinical performance examination. *Acad Med.* 2005;80:S34-7.
58. **Spickard A, Gabbe SG, Christensen JF.** Mid-career burnout in generalist and specialist physicians. *JAMA.* 2009;288:1447-50.
59. **Asch DA, Jedrzejewski MK, Christakis NA.** Response rates to mail surveys published in medical journals. *J Clin Epidemiol.* 1997;50:1129-36.
60. **Kellerman SE, Herold J.** Physician response to surveys. A review of the literature. *Am J Prev Med.* 2001;20:61-7.