

Ability of the Physician Well-Being Index to Identify Residents in Distress

LISELOTTE N. DYRBYE, MD, MHPE
 DANIEL SATELE, BA
 JEFF SLOAN, PhD
 TAIT D. SHANAFELT, MD

Abstract

Background Resident well-being impacts competence, professionalism, career satisfaction, and the quality of care delivered to patients.

Objective We established normative scores and reported evidence of relationship between the Physician Well-Being Index (PWBI) score to other variables and consequence validity for the PWBI in a national sample of residents, and evaluated the performance of the index after substituting the original fatigue item with an item not associated with driving a car.

Methods We conducted a cross-sectional survey study of a national sample of 20 475 residents. The survey included the PWBI, instruments assessing mental quality of life (QOL) and fatigue, and items on recent suicidal ideation and medical error. Fisher exact test or Wilcoxon/

2-sample *t* test procedures were used with a 5% type I error rate and a 2-sided alternative.

Results Of 7560 residents who opened the e-mail to participate in the study, 1701 (22.5%) completed the survey. Residents with low mental QOL, high fatigue, or recent suicidal ideation were more likely to endorse each of the PWBI items and a greater number of total items (all $P < .001$). At a threshold score of ≥ 5 , the PWBI's specificity for identifying residents with low mental QOL, high fatigue, or recent suicidal ideation was 83.6%. PWBI score also stratified residents' self-reported medical errors. The PWBI performed similarly using either fatigue item.

Conclusions The 7-item PWBI appears to be a useful screening index to identify residents whose degree of distress may negatively impact the quality of care they deliver.

Editor's Note: The online version of this article contains details about the Physician Well-Being Index used in this study.

Introduction

Residents' mental health can impair their competency, professionalism, career satisfaction, and the quality of care delivered to patients.¹⁻⁶ Burnout and other types of distress increase the risk of alcohol abuse, suicidal ideation, and

motor vehicle accidents.⁷⁻¹⁰ Unfortunately, many trainees hesitate to seek help and some opt to self-prescribe antidepressants,^{9,11-13} which is a risky and inappropriate strategy.¹⁴⁻¹⁷ Barriers to help seeking may include self-doubt about the need for treatment, perceived stigma, concerns about negative consequences, and a professional culture of stoicism.^{9,18,19}

Existing tools to evaluate distress are long, have complex scoring, and typically measure only 1 domain of distress (eg, fatigue, burnout). We first developed the Medical Student Well-Being Index (MSWBI), a 7-item instrument to screen for multiple dimensions of distress that commonly affect medical students. Evidence of content-related validity, internal structure validity, and reliability of the MSWBI, as well as how the MSWBI score relates to 3 clinically relevant outcomes (low mental quality of life [QOL], suicidal ideation, or serious thoughts of dropping out of medical school) and methods for establishing cut-scores in a national sample of medical students has been published.^{20,21} A slightly modified version, the Physician Well-Being Index (PWBI), was subsequently developed and tested in a national sample of 6994 practicing physicians in 2011.²² That study provided further validity evidence by showing the PWBI score not only stratifies mental QOL and recent suicidal ideation, but

All authors are at the Mayo Clinic. **Liselotte N. Dyrbye, MD, MHPE**, is Associate Professor of Medicine, Division of Primary Care Internal Medicine, Department of Medicine, College of Medicine; **Daniel Satele, BA**, is Statistician in Biomedical Statistics and Informatics, Department of Health Sciences Research; **Jeff Sloan, PhD**, is Professor of Biostatistics and Oncology, Department of Health Sciences Research; and **Tait D. Shanafelt, MD**, is Professor of Medicine, Division of Hematology, Department of Medicine, College of Medicine.

Funding: Funding for this study was provided by the American Medical Association and the Mayo Clinic Department of Medicine Program on Physician Well-Being.

Corresponding author: Liselotte N. Dyrbye, MD, MHPE, 200 First Street SW, Rochester, MN 55905, 507.284.2511, 507.266.2297, dyrbye.liselotte@mayo.edu

Received March 22, 2013; revision received August 5, 2013; accepted September 23, 2013.

DOI: <http://dx.doi.org/10.4300/JGME-D-13-00117.1>

also degree of fatigue, career satisfaction, and important clinical practice measures (intent to leave current practice and self-reported medical error) within the context of a diverse sample of physicians out in practice and done with training.

In the present study, we extend this work by gathering validity evidence within the context of a national sample of residents as the working conditions, work-hour expectations, and distress levels of residents are distinct from both medical students and practicing physicians. We report evidence of relationship between index score to other variables and consequence validity for the PWBI in this national sample of residents. We also evaluate the performance of the index after substituting the original fatigue item with a new fatigue item not dependent on a need to regularly drive a car.

Methods

Participants

As previously reported,²³ in 2012 we surveyed 20 475 US residents/fellows (“residents”) listed in the Physician Masterfile (PMF), which contains nearly all residents independent of American Medical Association (AMA) membership, who had an e-mail address on file with the AMA, and permitted it to be used for correspondence. Invitations were sent to these residents; 7560 residents opened at least 1 e-mail invitation. We considered those who opened an e-mail to have received an invitation to participate in the study.²⁴ Participation was voluntary and all responses were anonymous.

Study Measures

The survey included the PWBI and items inquiring about demographics, specialty area, recent suicidal ideation, medical error, and standardized instruments to measure QOL.

Physician Well-Being Index

We have previously reported the methods used to develop the 7-item PWBI. Details are provided as online supplemental material. Briefly, the index is intended to include the domains of burnout, depression, stress, fatigue, and mental and physical QOL. It consists of 7 yes/no items and respondents receive a score from 0 to 7 based on responses.²¹ Our previous studies suggest a threshold score of ≥ 4 for medical students and practicing physicians.^{20,22} At a threshold score of ≥ 4 , the specificity for detecting medical students and practicing physicians with low mental QOL was 87.7% and 81.0%, respectively, and the sensitivity was 59.2% and 73.3%, respectively.

Fatigue

As not all residents drive cars regularly we explored whether the performance of the PWBI improved when an

What was known

Earlier work shows some residents are in distress during training, with a negative impact on competence, career satisfaction, and quality of care.

What is new

A 7-item instrument identified residents in distress, with a high distress score also associated with a greater number of residents' self-reported medical errors.

Limitations

Survey study assessed association, but not the direction of the relationship or causality; low response rate and the potential for respondent bias.

Bottom line

A brief screening tool with adequate construct validity may be used to identify residents who may benefit from added resources, or in resident self-assessment and subsequent help seeking.

alternative fatigue item not dependent on a need to drive (“have you fallen asleep while sitting inactive in a public place”) was used in place of the original PWBI fatigue item (“have you fallen asleep while stopped in traffic or driving”). Both items stemmed from the Epworth Sleepiness Scale,²⁵ which was identified during our original instrument development phase to capture the intended construct.²⁶ All participants were asked both fatigue items and index scores were calculated for each participant using both the original and alternative fatigue item to see if 1 item improved performance.

Other Study Measures

Residents rated their mental QOL and level of fatigue over the past week on a standardized linear analog scale. This scale has evidence of validity in a variety of medical conditions and populations^{27,28} and has been used in other studies of residents.⁵ The item inquiring about recent suicidal ideation within the last 12 months is similar to questions used in large US epidemiologic studies^{29,30} and has been used in previous samples of physicians.⁹ One item identical to that used in a previous study of residents and practicing physicians inquired about self-perceived medical errors in the last 3 months.^{5,31}

Relationship to Other Variables

As distress has multiple dimensions (eg, depression, burnout, fatigue, QOL) there is not a single gold standard for defining “severe distress.” In this study, we assessed the ability of the PWBI to:

1. Identify residents with low mental QOL defined by a score $\geq 1/2$ SD *below* the sex-matched general population norm⁵ (a clinically meaningful effect size³²)
2. Identify residents who had high level of fatigue defined by a score $\geq 1/2$ SD *below* the sex-matched

general population norm (lower scores indicate higher fatigue)³³

3. Identify residents who reported suicidal ideation within the last 12 months as an alternative measure of distress because it represents a clinically relevant outcome that warrants individualized counseling
4. Stratify residents' likelihood of reporting a recent major medical error to identify those in need of support that could reduce inappropriate self-blame

The Mayo Clinic Institutional Review Board approved the study.

Statistical Analysis

We used basic descriptive statistics and Fisher exact test or Wilcoxon/2-sample *t* test procedures, as appropriate. We used a 5% type I error rate and a 2-sided alternative. We calculated the sensitivity, specificity, and likelihood ratios (LRs) associated with PWBI scores for outcomes of interest. Lastly, we used a hypothetical cohort to explore the practicality of using the PWBI at various cut points. We conducted all analysis using SAS version 9 (SAS Institute Inc, Cary, NC).

Results

Of the 7560 residents who opened the e-mail invitation, 1701 residents (22.5% participation rate) completed the survey. The demographics of responders in comparison to the 12 9608 residents listed in the PMF were generally similar, although fewer responders were male (participants 49% versus PMF 54%).²³ The specialty distribution of responders was similar to US residents in general,³⁴ with the exception of slightly fewer responders in family medicine (6.9% versus 10.0%) and a larger proportion in obstetrics-gynecology (11.0% versus 5.2%). Rates of burnout, symptoms of depression, recent suicidal ideation, high fatigue, and QOL have been previously reported²³ and were similar to findings in other studies with substantially higher participation rates.²

Mental QOL

Residents with low mental QOL were more likely to endorse each PWBI item (TABLE 1) as well as a greater total number of items (mean 5.0 [SD 1.5] versus mean 2.7 [SD 2.0], $P < .001$). As the number of PWBI items endorsed increased so did the odds of having low mental QOL. The likelihood ratio for low mental QOL among residents with PWBI scores ≥ 5 was 3.42 in comparison to 0.37 for those with scores < 5 .

Using exact PWBI scores, which provide a way to estimate an *individual* resident's risk of distress after they complete the PWBI, the LR of low mental QOL ranges from 0.06 to 5.66 (TABLE 2). Assuming a 28.9%

prevalence of low mental QOL (ie, the approximate prevalence in the overall 2012 AMA sample²³ and in previous large studies of physicians⁶) as the pretest probability, the PWBI exact score can lower the posttest probability to 2.4%, or raise it to 70.0%.

Fatigue and Suicidal Ideation

Residents with high levels of fatigue or recent suicidal ideation were more likely to endorse each PWBI item and a greater number of total items (all $P < .001$). As the number of PWBI items endorsed increased so did the odds of high fatigue (odds ratio [OR] 0.34–5.189) and suicidal ideation (OR 0.22–4.65). Assuming a prevalence of 32.5% for high fatigue⁵ as the pretest probability, the PWBI exact score can lower the posttest probability to 15.4% or raise it to 68.7% (TABLE 2). Similarly, using a prevalence of 7.98% for recent suicidal ideation^{9,23} as the pretest probability, the PWBI exact score can lower the posttest probability to $< 1\%$, or raise it to 25%.

For each outcome (ie, low mental QOL, high level of fatigue, and recent suicidal ideation), the likelihood ratio was less than 1 for those with an exact score of 4 or below and greater than 1 for those with exact scores of 5, 6, or 7, suggesting 5 is the inflection point.

Threshold Scores

Threshold scores (ie, ≥ 1 etc) can estimate the risk of distress in a group scoring at or above a specific threshold. Thus, they can be useful for establishing a cutoff score to identify a subset of residents who may benefit from further evaluation or support. At a threshold score of ≥ 5 , the specificity of the PWBI for identifying residents with low mental QOL was 79.4% and the sensitivity was 70.3%.

Next, we examined the prevalence of high fatigue or suicidal ideation among residents who had index scores ≥ 5 but did not have low mental QOL. These residents would be considered “false positives” based on their mental QOL score alone. Among the 244 residents with index scores ≥ 5 who did not have low mental QOL, 84 had high fatigue and 27 reported recent suicidal ideation suggesting that they were experiencing substantial distress despite their mental QOL scores. When only those residents with PWBI scores ≥ 5 who did not have low mental QOL, high fatigue, and/or recent suicidal ideation were considered false positives, the specificity increased to 83.6%.

PWBI Scores and Practice-Related Outcome

Residents who reported a recent major medical error were also more likely to endorse each of the individual PWBI items and a greater number of total items (both $P < .001$). Assuming an estimated prevalence of 16.3% for recent major medical error^{5,23} as the pretest probability, the PWBI

TABLE 1 **PHYSICIAN WELL-BEING INDEX (PWBI) ITEMS ENDORSED BY 1667 US RESIDENTS WITH AND WITHOUT LOW MENTAL QUALITY OF LIFE (QOL), 2012^a**

Item	No. (%) Endorsing Item ^b		OR (95% CI) ^c
	Residents With Low Mental QOL (n = 481)	Residents Without Low Mental QOL (n = 1186)	
During the past month:			
1. Have you felt burned out from your work?	451 (93.8)	717 (60.5)	9.83 (6.674, 14.49)
2. Have you worried that your work is hardening you emotionally?	400 (83.2)	552 (46.5)	5.67 (4.353, 7.39)
3. Have you often been bothered by feeling down, depressed, or hopeless?	375 (78.0)	360 (30.4)	8.12 (6.330, 10.41)
4. Have you fallen asleep while stopped in traffic or driving?	167 (34.7)	205 (17.3)	2.546 (2.00, 3.24)
5. Have you felt that all the things you had to do were piling up so high that you could not overcome them?	367 (76.3)	536 (45.2)	3.90 (3.073, 4.96)
6. Have you been bothered by emotional problems (such as feeling anxious, depressed, or irritable)?	439 (91.3)	640 (54.0)	8.92 (6.369, 12.49)
7. Has your physical health interfered with your ability to do your daily work at home and/or away from home?	229 (47.6)	190 (16.0)	4.76 (3.759, 6.04)
No. of Items Endorsed			
0	5 (1.0)	203 (17.1)	0.05 (0.02, 0.12)
1	10 (2.1)	189 (15.9)	0.11 (0.06, 0.21)
2	23 (4.8)	196 (16.5)	0.25 (0.16, 0.40)
3	29 (6.0)	160 (13.5)	0.41 (0.27, 0.62)
4	76 (15.8)	194 (16.4)	0.96 (0.72, 1.28)
5	125 (26.0)	137 (11.6)	2.69 (2.05, 3.52)
6	135 (28.1)	73 (6.2)	5.95 (4.37, 8.10)
7	78 (16.2)	34 (2.9)	6.56 (4.32, 9.97)

^a For the purposes of this study, low mental QOL was defined as having a mental standardized linear analog QOL score $\geq 1/2$ standard deviation below that of the sex-matched general population.

^b All PWBI questions are answered "yes" or "no." One point is assigned for each "yes" response.

^c Odds ratio represents risk of low mental QOL in group of physicians who endorsed the item or the number of items relative to the referent group. The PWBI is copyrighted, and permission for use must be obtained by the author.

exact score can lower the posttest probability to 4.8% or raise it to 39.3% (TABLE 2).

Alternative Fatigue Item

Overall, the PWBI performed similarly using either fatigue item. For example, the posttest probability for high fatigue using the alternative fatigue item ranged from 16.5% (score of 0) to 65.4% (score of 7), and sensitivity and specificity for high fatigue differed by only 2.0% to 3.0%.

Use of the PWBI to Screen a Hypothetical Cohort of Residents

Lastly, we modeled the outcome of screening a hypothetical cohort of 100 residents using the PWBI and a PWBI score ≥ 5 to identify residents at risk. Assuming a prevalence of

low mental QOL, fatigue, and recent suicidal ideation similar to that of the 2012 national sample of US residents, 35 would have a score ≥ 5 . Of these 35 residents, 26 would have low mental QOL, recent suicidal ideation, or high fatigue.

Discussion

In this cohort of 1701 residents representing all specialties and training stages, the PWBI stratified residents' well-being and identified those with low mental QOL, high fatigue, and recent suicidal ideation. At a threshold score of ≥ 5 , the PWBI's specificity for identifying residents with low mental QOL is 79.4% and the sensitivity is 70.3%. The specificity improved to 83.6% when we defined

TABLE 2 EFFICACY OF THE PHYSICIAN WELL-BEING INDEX (PWBI) FOR IDENTIFYING LOW MENTAL QUALITY OF LIFE (QOL), UNFAVORABLE FATIGUE QOL, RECENT SUICIDAL IDEATION, AND SELF-PERCEIVED MEDICAL ERROR AMONG 1667 US RESIDENTS^a

PWBI Exact Score ^b	Low Mental QOL (n = 481)		High Fatigue (n = 541)		Suicidal Ideation (n = 133)		Self-Perceived Recent Medical Error (n = 269)	
	LR ^c	Prob. ^d %	LR	Prob. %	LR	Prob. %	LR	Prob. %
0	0.06	2.4	0.38	15.4	0	0.0	0.26	4.8
1	0.13	5.0	0.29	12.1	0.24	2.0	0.36	6.6
2	0.29	10.5	0.67	24.2	0.16	1.4	0.73	12.5
3	0.45	15.3	0.71	25.4	0.52	4.3	0.66	11.5
4	0.97	28.1	0.68	24.4	0.78	6.3	0.92	15.2
5	2.25	47.7	1.88	47.3	1.34	10.3	1.81	26.1
6	4.56	64.9	2.63	55.8	3.21	21.6	1.47	22.3
7	5.66	69.6	4.59	68.7	3.87	25.0	3.31	39.3

^a We defined (1) low mental QOL as having a mental standardized linear analog QOL score $\geq 1/2$ SD below that of the sex-matched general population; (2) high fatigue as having a fatigue standardized linear analog score $\geq 1/2$ SD below that of the sex-matched general population (high score is favorable); and (3) recent suicidal ideation as endorsing experiencing suicidal ideation within the previous 12 months.

^b The PWBI exact score is the number of the 7 PWBI items endorsed.

^c LR indicates the likelihood ratio associated with the exact score.

^d Posttest probability was calculated using an estimated prevalence of 28.85% for low mental QOL, 32.45% for high fatigue, 7.98% for suicidal ideation, and 16.3% for medical error as the pretest probability.

false positives as not having low mental QOL, high fatigue, or recent suicidal ideation—a reasonable approach as having these outcomes warrant identification and response regardless of a mental QOL score. At this threshold score the PWBI's specificity is comparable to that of other established and widely used screening instruments.^{35–40} Notably, the threshold score of ≥ 5 is different from the threshold score previously identified in large samples of medical students and practicing physicians.^{20,22} This reinforces the need to evaluate the PWBI in a large sample of residents as was done in this study. We also found that replacing the original fatigue item with a new fatigue item not dependent on a need to drive did not substantially alter the performance of the PWBI.

Favorable characteristics of the PWBI include its brevity, ease of administration, simple scoring strategy, breadth of dimensions of distress covered, and ability to identify residents whose degree of distress places them at risk for committing a medical error. Hence, it is well-suited for a brief self-assessment tool that residents could regularly use to assess their current level of distress, and gain insight into when their distress is placing them at risk for potentially serious consequences. Residents who completed the PWBI could also receive immediate personalized feedback about how their level of distress compares to

peers, and information about local mental health and wellness resources. Thus, it has the potential to help residents self-monitor, promote their own wellness, and recognize when they need help.

Residency programs could use the PWBI as part of a screening process to improve resource allocation to those residents in greatest need of support. Such a proactive approach has the potential to lead to earlier intervention when a resident's distress may be less severe, and before it has led to adverse consequences. Screening approaches may need to be pursued through an employee assistance program or a provost who is not connected to the residency program to ensure confidentiality is maintained. Based on this study's results a threshold score of 5 or greater appears reasonable. Alternatively, aggregate, deidentified resident scores on the PWBI could also be funneled to residency programs to provide global information in regards to trainee well-being.

This study has several limitations. First, distress is a multidimensional construct and no gold standard exists for measuring it. We examined 3 clinically relevant dimensions of distress (low mental QOL, high fatigue, and suicidal ideation) with potential for serious personal and professional consequences. The PWBI may not stratify risk for other dimensions of distress. However, as mean mental QOL decreased in a stepwise fashion with each 1 point

increase in PWBI score the PWBI appears to be a powerful risk stratification tool. Second, the PWBI is a screening tool and not a diagnostic instrument. Hence, it is intended to improve residents' self-awareness, provide calibration relative to peers, and identify those who may benefit from further evaluation or support. Third, we conducted a cross-sectional study, with the potential for selection bias. Future studies are needed to evaluate the consequences of residents completing the PWBI (including subsequent allocation of resources or help seeking), as well as the instrument's ability to assess change over time. Strengths include our methodological approach of evaluating the instrument in a national sample of residents while simultaneously using well-established metrics and identifying residents whose degree of distress is associated with clinically important outcomes and relevant practice-related risks.

Conclusion

Our results offer further evidence of construct validity of the PWBI in a selected national sample of residents. Further research is needed to determine the utility of using the PWBI as a screen to identify residents who may benefit from individual resources or for residents to self-assess their level of distress, and whether screening with the PWBI facilitates help-seeking behaviors.

References

- Dyrbye LN, Massie FS Jr, Eacker A, Harper W, Power D, Durning SJ, et al. Relationship between burnout and professional conduct and attitudes among us medical students. *JAMA*. 2010;304(11):1173–1180.
- West CP, Shanafelt TD, Kolars JC. Quality of life, burnout, educational debt, and medical knowledge among internal medicine residents. *JAMA*. 2011;306(9):952–960.
- West CP, Huschka MM, Novotny PJ, Sloan JA, Kolars JC, Habermann TM, et al. Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. *JAMA*. 2006;296(9):1071–1078.
- West CP, Shanafelt TD. The influence of personal and environment factors on professionalism in medical education. *BMC Med Educ*. 2007;7.
- West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD. Association of resident fatigue and distress with perceived medical errors. *JAMA*. 2009;302(12):1294–1300.
- Shanafelt TD, Balch CM, Bechamps GJ, Russell T, Dyrbye L, Satele D, et al. Burnout and career satisfaction among American surgeons. *Ann Surg*. 2009;250(3):463–471.
- Oreskovich MR, Kaups KA, Balch CM, Hanks JB, Satele D, Sloan J, et al. The prevalence of alcohol use disorders among american surgeons. *Arch Surg*. 2011;147(2):168–174.
- Dyrbye LN, Thomas MR, Massie FS, Power DV, Eacker A, Harper W, et al. Burnout and suicidal ideation among US medical students. *Ann Intern Med*. 2008;149(5):334–341.
- Shanafelt TD, Balch CM, Dyrbye L, Bechamps G, Russell T, Satele D, et al. Special report: suicidal ideation among American surgeons. *Arch Surg*. 2011;146(1):54–62.
- West CP, Tan AD, Shanafelt T. Association of resident fatigue and distress with occupational blood and body fluid exposures and motor vehicle incidents. *Mayo Clin Proc*. 2012;87(12):1138–1144.
- Montgomery AJ, Bradley C, Rochfort A, Panagopoulou E. A review of self-medication and medical students. *Occup Med (Lond)*. 2011;61(7):490–497.
- Christie JD, Rosen IM, Bellini LM, Inglesby TV, Lindsay J, Alper A, et al. Prescription drug use and self-prescription among resident physicians. *JAMA*. 1998;280(14):1253–1255.
- Guille C, Sen S. Prescription drug use and self-prescription among training physicians. *Arch Intern Med*. 2012;172(4):371–372.
- American Medical Association. AMA's Code of Medical Ethics. Self-Treatment or Treatment of Immediate Family Members. <http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/code-medical-ethics/opinion819.page?>. Accessed November 4, 2013.
- Snyder L; American College of Physicians Ethics, Professionalism, and Human Rights Committee. American College of Physicians Ethics Manual: sixth edition. *Ann Intern Med*. 2012;156(1, pt 2):73–104.
- American Psychiatric Association. Opinions of the Ethics committee on the principles of medical ethics. 2009. <http://www.psychiatry.org/practice/ethics/ethics>. Accessed January 23, 2013.
- National Alliance for Physician Competence. Guide to Good Medical Practice—USA. 2008. <http://www.ama-assn.org/ama1/pub/upload/mm/377/ggmp-usa.pdf>. Accessed November 4, 2013.
- Givens JL, Tjia J. Depressed medical students' use of mental health services and barriers to use. *Acad Med*. 2002;77(9):918–921.
- Center C, Davis M, Detre T, Ford DE, Hansbrough W, Hendin H, et al. Confronting depression and suicide in physicians: a consensus statement. *JAMA*. 2003;289(23):3161–3166.
- Dyrbye LN, Schwartz A, Downing SM, Szydlo DW, Sloan J, Shanafelt T. Efficacy of a brief screening tool to identify medical students in distress. *Acad Med*. 2011;86(7):907–914.
- Dyrbye LN, Szydlo DW, Schwartz A, Downing SM, Sloan J, Shanafelt T. Development and preliminary psychometric properties of a well-being index for medical students. *BMC Med Educ*. 2010;10.
- Dyrbye LN, Satele D, Sloan J, Shanafelt T. Utility of a brief screening tool to identify physicians in distress. *J Gen Intern Med*. 2013;28(3):421–427.
- Dyrbye LN, West CP, Satele D, Boone S, Tan L, Sloan J, et al. Burnout among US medical students, residents, and early career physicians relative to the general US population. *Acad Med*. 2014 Jan 20. [Epub ahead of print]
- American Association for Public Opinion Research. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Revised 2011. http://www.aapor.org/AM/Template.cfm?Section=Standard_Definitions&Template=/CM/ContentDisplay.cfm&ContentID=3156. Accessed November 4, 2013.
- Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540–545.
- Dyrbye LN, Szydlo DW, Downing SM, Sloan JA, Shanafelt TD. Development and preliminary psychometric properties of a well-being index for medical students. *BMC Med Educ*. 2010;10:8.
- Gudex C, Dolan P, Kind P, Williams A. Health state valuations from the general public using the visual analogue scale. *Qual Life Res*. 1996;5(6):521–531.
- Rumman TA, Clark MM, Sloan JA, Frost MH, Bostwick JM, Atherton PJ, 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(4):635–642.
- National Comorbidity Survey. Collaborative Psychiatric Epidemiology Surveys. Suicidality. <http://www.icpsr.umich.edu/cocoon/cpes/cpes/BLSUICIDALITY/all/section.xml>. Accessed May 15, 2008.
- United States Department of Health and Human Services. Youth Risk Behavior Survey. Centers for Disease Control and Prevention. National Center for Health Statistics. <http://www.cdc.gov/HealthyYouth/yrbs/trends.htm>. Accessed March 24, 2010.
- Shanafelt TD, Balch CM, Bechamps G, Russell T, Dyrbye L, Satele D, et al. Burnout and medical errors among American surgeons. *Ann Surg*. 2010;251(6):995–1000.
- Norman GR, Sloan JA, Wyrwich KW. The truly remarkable universality of half a standard deviation: confirmation through another look. *Expert Rev Pharmacoecon Outcomes Res*. 2004;4(5):581–585.
- Shanafelt TD, Boone S, Tan L, Dyrbye LN, Sotile W, Satele D, et al. Burnout and satisfaction with work-life balance among US physicians relative to the general US population. *Arch Intern Med*. 2012;172(18):1377–1385.
- Accreditation Council for Graduate Medical Education. Data Resource Book. Academic Year 2010–2011. http://www.acgme.org/acgmeweb/Portals/0/PFAssets/PublicationsBooks/2010-2011_ACGME_DATA_RESOURCE_BOOK.pdf. Accessed April 17, 2013.
- Spitzer RL, Williams JB, Kroenke K, Linzer M, deGruy FV 3rd, Hahn SR, et al. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA*. 1994;272(22):1749–1756.
- 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(7):439–445.

- 37 Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606–613.
- 38 Hirschfeld RM, Williams JB, Spitzer RL, Calabrese JR, Flynn L, Keck PE Jr, et al. Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. *Am J Psychiatry.* 2000;157(11):1873–1875.
- 39 Kroenke K, Spitzer RL, Williams JB, Monahan PO, Lowe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med.* 2007;146(5):317–325.
- 40 Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092–1097.