Longitudinal Relationships Between Fatigue and Depression In Cancer Patients with Depression and/or Pain

Linda F. Brown, Kevin L. Rand, Silvia M. Bigatti, Jesse C. Stewart, Dale E. Theobald, Jingwei Wu, and Kurt Kroenke

Abstract

Objective—Fatigue is one of the most common and debilitating symptoms reported by cancer patients, yet relatively little is understood about its etiology. Recently, as researchers have begun to focus attention on cancer-related fatigue (CRF), depression has emerged as its strongest correlate. Few longitudinal studies, however, have examined directionality of the relationship between the two symptoms. Our aim was to evaluate the directionality of the association between depressive symptoms and fatigue.

Methods—The study used a single-group cohort design of longitudinal data (N = 329) from a randomized controlled trial of an intervention for pain and depression in a heterogeneous sample of cancer patients. Participants met criteria for clinically significant pain and/or depression. Our hypothesis that depression would predict change in fatigue over 3 months was tested using latent variable cross-lagged panel analysis.

Results—Depressive symptoms and fatigue were strongly correlated in the sample (baseline correlation of latent variables = 0.71). Although the model showed good fit to the data, χ² (66, N = 329) = 88.16, p = 0.04, SRMR = 0.030, RMSEA = 0.032, and CFI = 1.00, neither structural path linking depression and fatigue was significant, suggesting neither symptom preceded and predicted the other.

Conclusions—Our findings did not support hypotheses regarding the directionality of the relationship between depressive symptoms and fatigue. The clinical implication is that depression-specific treatments may not be sufficient to treat CRF and that instead, interventions specifically targeting fatigue are needed.
Fatigue is a vexing problem for individuals with cancer. It is the most common symptom reported by cancer patients (Berger et al., 2010), adds considerably to suffering, and exists across all disease types and stages. It has been found to be a problem before, during, and after treatment, sometimes continuing long after treatment has ended, even in those believed to be free of disease (Hofman, Ryan, Figueroa-Moseley, Jean-Pierre, & Morrow, 2007). Cancer-related fatigue (CRF) has been reported by up to 40% of patients at diagnosis, 90% of patients receiving radiation treatment, and 80% of those undergoing chemotherapy (Hofman et al., 2007). In view of its prevalence and detrimental impact on quality of life, CRF is an important symptom to target in treatment, yet it is too seldom diagnosed or treated (Higginson, Armes, & Krishnasamy, 2004). Moreover, there is still a paucity of evidence-based treatments for CRF.

CRF is a complex and multidimensional symptom that is known to occur both as a consequence of the cancer itself and as a side effect of treatment (Hofman et al., 2007; Mustian et al., 2007; Ryan et al., 2007). It has recently gained more attention as researchers and clinicians seek to improve cancer symptom management (Patrick et al., 2002). CRF has been defined as “a distressing, persistent, subjective sense of physical, emotional, and/or cognitive tiredness or exhaustion related to cancer or cancer treatment that is not proportional to recent activity and interferes with usual functioning” (Berger et al., 2010, p. FT-1).

Although the etiology of CRF is poorly understood, it has consistently been found to have a strong association with depression (Brown & Kroenke, 2009). In fact, reviews of studies of the relationship of depression to fatigue have found the magnitude of the association to be the highest among all the correlates of fatigue, including disease activity as measured by such markers as tumor-specific tests and nutritional status (Hotopf, 2004). Whether a directional relationship that might suggest causality exists, however, is unknown. The current study was undertaken to extend the understanding of the relationship of depression and CRF in a way that will inform the development of effective treatment of these symptoms. It is important to know whether reductions in depression lead to decreased CRF, if the converse is true (i.e., whether reductions in CRF lead to decreased depression), and to what degree effects are bidirectional. A predominant effect in one direction (i.e., depression improvement accounting for much of the improvement in fatigue) might support a “treat depression first” strategy. In contrast, a bidirectional relationship might suggest interventions that either treat fatigue and depression as a cluster (i.e., treatments that have proven effective for both symptoms) or that comprise “dual-diagnosis” treatments (i.e., fatigue-specific therapy coupled with depression-specific therapy).

Longitudinal studies with published findings about the directionality of the relationship between fatigue and depression in cancer are few, and results are mixed and inconclusive. In only 2 studies have authors clearly asserted that their findings suggest that changes in fatigue were associated with changes in depression (Stone, Richards, A’Hern, & Hardy,
Authors of at least 6 studies concluded that no evidence had been found in their research to suggest that depression was related to longitudinal changes in fatigue (Geinitz et al., 2001; Morrow et al., 2003; Pirl, 2008; Schumacher et al., 2002; Stone, Hardy, Huddart, A., & Richards, 2000; Visser & Smets, 1998). None of the published studies employed panel analysis, however, which is suitable for examining hypotheses of directionality or causality because it can provide evidence regarding all three conditions of causality: 1) covariation of the 2 variables; 2) time precedence of the putative causal variable; 3) and nonspuriousness (i.e., the association of the 2 variables must not be produced by a joint association with a third variable or set of variables; Finkel, 1995).

The current study was undertaken to add to what is known about the interrelationship between CRF and depression by using a sophisticated statistical approach to the analysis. We used latent variable analysis as a way of accounting for measurement error (Shadish, Cook, & Campbell, 2002). Our study examined the relationship of depression and fatigue over time in a heterogeneous sample of cancer patients. It was a secondary analysis of data from a randomized controlled trial of an intervention to treat depression and pain.

Hypotheses were derived from a perpetuating factors model of fatigue —developed from what was available in extant literature— that suggested that depression is one of several emotional, cognitive, and behavioral factors that perpetuate fatigue (Hotopf, 2004). Our first hypothesis was that initial levels of depression would predict longitudinal changes in fatigue. Our second hypothesis was that initial levels of fatigue would also predict future changes in depression—but at a lower magnitude. If depression was supported as a predictor (i.e., perpetuator) of CRF, then interventions targeting depressive symptoms may be appropriate in treating fatigued cancer patients.

**Methods**

**Research Design**

This study used a single-group cohort design in a secondary analysis of longitudinal data from a randomized controlled trial (N = 405) of an intervention for pain and depression in cancer patients. The Indiana Cancer Pain and Depression trial (INCPAD) tested the effectiveness of telecare management delivered by a nurse-psychiatrist team in a statewide network of urban and rural community-based cancer clinics. The intervention was based upon the empirically-validated Three-Component Model (TCM; Dietrich et al., 2004; Oxman, Dietrich, Williams Jr., & Kroenke, 2002). In INCPAD, the model was a collaboration between the oncology practice, a centralized nurse care manager, and a supervising pain-psychiatrist. A telemedicine approach was utilized with automated home-based monitoring of pain and depressive symptomatology coupled with telephonic nurse care management over 12 months. Medication management utilized evidence-based algorithms for antidepressants and analgesics. Participants were assessed at baseline, 1, 3, 6, and 12 months by telephone interviewers blinded to treatment group. The sample included patients with cancer-related pain (n = 96), clinical depression (n = 131), or both (n = 178). Participants were randomized by computer to either the intervention or usual care control group, stratified by symptom type. Participants were enrolled from March 2006 to August...
Details of the study design and outcomes have been published (Kroenke et al., 2009; Kroenke, Theobald et al., 2010). Patients in the intervention group had greater improvements than the usual care control group in both depression and pain at all time points.

The present study uses data from the baseline and 3-month assessments for several reasons. First, the most intense phase of the INCPAD depression intervention occurred during the first three months of the 12-month trial. Second, depression improvement was maximal at 3 months after which it largely plateaued. Third, data at 3 months are less affected by attrition due to death, drop-out, or loss to follow-up than at later time-points. At 3 months, the intervention in the INCPAD RCT produced a moderate improvement in depression (effect size of 0.42, \( p < .001 \)). A small improvement in fatigue was also found (effect size of 0.20, \( p = .03 \)).

**Participants**—Participants were recruited from 16 oncology clinics. Participant characteristics are presented in Table 1. The sample was predominantly Caucasian and female, and the average age was 58 years (SD = 10.8). Although the INCPAD sample was recruited based on clinically significant pain or depression, fatigue was prevalent at baseline. In fact, in a published secondary analysis of somatic symptom burden in INCPAD, fatigue was reported to be the most bothersome symptom among the 22 symptoms assessed, with 98% of the sample reporting feeling tired and 79% reporting being “bothered a lot” by this symptom (Kroenke, Zhong et al., 2010). Moreover, the mean score on the SF vitality scale was 28.3 (SD = 19.2), which suggested clinically significant fatigue according to established cutoffs (Donovan, Jacobsen, Small, Munster, & Andrykowski, 2008).

**Sampling Procedures**

To identify eligible participants for the INCPAD study, oncology clinic patients completed a 4-item depression and pain questionnaire, and those who screened positive for pain or depression and expressed interest were contacted by telephone for eligibility and enrollment procedures. To be eligible, cancer patients met study criteria for pain or depression or both. Patients met criteria for depression if they had a PHQ-9 score of 10 or greater with endorsement of depressed mood and/or anhedonia. PHQ-9 scores range from 0 to 27 with higher scores representing more severe depression; a cutoff point of 10 indicates depression of at least moderate severity. Patients met criteria for pain if they had a score of 5 or greater on the Brief Pain Inventory, which is scored 0 to 10 with higher scores representing more severe pain (Cleeland, 1991); a cutoff point of 5 indicates pain of at least moderate severity. Additional eligibility criteria were that the pain had persisted after use of at least two different analgesics and was cancer-related. Individuals were excluded if they did not speak English, had moderately severe cognitive impairment as defined by a validated 6-item cognitive screener (Callahan, Unverzagt, Hui, Perkins, & Hendrie, 2002), had schizophrenia or other psychosis, had a pending pain-related disability claim, were pregnant, or were in hospice care.

During the 130-week enrollment period, 4,465 patients were screened in the 16 participating clinics. Of those, 1,851 screened positive for pain and/or depression and 616 were found to
be eligible for the INCPAD trial. Of the 405 participants enrolled, 131 (32.3%) had depression only, 96 (23.7%) pain only, and 178 (43.9%) both depression and pain.

**Measures**—Fatigue was measured with the vitality subscale of the SF-36 Health Survey (version 2), an instrument that assesses health-related quality of life (Ware, Snow, Kosinski, & Gandek, 1993). The vitality subscale assesses energy level and fatigue as a way to capture differences in subjective well-being. Respondents are asked “How much of the time during the past 4 weeks “did you have a lot of energy?” “…have you felt full of life?” “…did you feel worn out?” and “…did you feel tired?” A 5-level response scale ranges from none of the time to all of the time. Norm-based vitality scale cores range from 0 to 100, with lower scores representing more severe fatigue. When analyzing the vitality scale as a dichotomous measure, previous researchers have categorized scores above the midpoint of 50 as representing well-being and scores below 50 as indicating disability due to fatigue (Bower et al., 2006). The 25th percentile has been established in the literature as a clinically significant indicator of impairment; that is, those scoring below the 25th percentile may be classified as fatigued. As an example, the cutoff was established at 45 in a sample of females in the U.S. (Donovan et al., 2008).

The SF-36 has well established internal consistency, reliability, content validity, construct validity, and criterion-related validity, having been tested in a variety of population samples (Wu & McSweeney, 2004). In baseline data from the INCPAD study, Cronbach's coefficient alpha for the vitality scale was 0.77. The vitality scale has been widely used to assess fatigue across a range of conditions. A recent review (O'connor, 2004) surveyed the number of citations in medical and psychology databases for commonly used measures of energy and fatigue and found the SF-36 vitality scale to be the most-cited, with 2,449 references.

Because the SF-36 vitality scale is a general measure of fatigue rather than cancerspecific, we tested its convergent validity in a subset of the INCPAD sample with a well-validated scale developed for cancer-related fatigue—The Fatigue Symptom Inventory (FSI; Hann et al., 1998). Both scales were administered to a consecutive subsample of individuals interviewed at 1 month (n = 68) and 6 months (n= 96). Comparing scores of three subscales of the FSI to the vitality scale score, we found that the 4-item vitality scale performed similarly to that of the 13-item FSI (Brown, Kroenke, Theobald, & Wu, 2011).

The Depression Subscale (SCL-20) of the Hopkins Symptom Checklist (SCL-90) was the primary outcome measure for depression in the INCPAD study and one of three depression scales used in the current study. The SCL-20 was chosen for its demonstrated sensitivity in detecting differences in depression severity between treatment groups in previous trials (Kroenke, West et al., 2001; Unutzer et al., 2002). It has been used as a primary outcome for depressive symptoms in several trials (e.g., Dietrich et al., 2004; Kroenke, Spitzer, Williams, & Lüwe, 2010). In 20 items, respondents rate how much distress was experienced over the past 4 weeks because of various symptoms such as “feeling lonely or blue,” “feeling no interest in things,” “trouble falling asleep,” and “thinking, speaking, and moving at a slower pace.” For each item, five response choices range from “not at all” to “extremely.” In INCPAD, Cronbach's coefficient alpha was 0.89.
The Patient Health Questionnaire, Depression (PHQ-9) is the depression module of the Patient Health Questionnaire, a self-administered diagnostic instrument for common mental disorders developed for use in primary care settings (Kroenke, Spitzer, & Williams, 2001). Its nine depression items comprise the nine criteria upon which the diagnosis of depressive disorders is based (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition Text Revision, 2000). In 2 studies including 6,000 patients, the PHQ-9 performed well as a brief measure of depression severity and showed good construct and criterion validity (Kroenke, Spitzer et al., 2001). It has been recognized as useful both for making diagnoses and assessing severity. Respondents are asked whether they have been bothered by a specified symptom over the last 2 weeks. Examples of symptoms are “little interest or pleasure in doing things,” and “feeling bad about yourself, feeling that you are a failure, or feeling that you have let yourself or your family down.” Participants are asked whether they were bothered “several days,” “more than half the days,” or “nearly every day.” Cronbach's coefficient alpha for the PHQ-9 in the INCPAD sample was 0.81.

The Mental Health Inventory (MHI-3) is a depression subscale of the Mental Health Inventory-5, which in turn is part of the SF-36 Health Survey. The MHI-3 recently has been validated as a measure for depression (Cuijpers, Smits, Donker, ten Have, & de Graeff, 2009; Rumpf, Meyer, Hapke, & John, 2001). The MHI-5 consists of 5 items from the SF-36 and assesses mental health (Rumpf et al., 2001). It has been tested in a large sample of the general population (n = 4,036) for its ability to screen for various mental disorders, especially depression and anxiety. It performed best in detecting depressive disorders, with sensitivity and specificity statistics comparable to 9 lengthier instruments used in primary care. More recently, Cuijpers and colleagues (2009) undertook ROC analysis and found no difference between the MHI-5 and the MHI-3 in detecting major depression and dysthymia. The MHI-3 items ask “How much of the time during the last month have you: 1.) felt downhearted and depressed?, 2.) been happy? and 3.) felt so down in the dumps that nothing could cheer you up?” Respondents choose from 5 options ranging from “none of the time” to “all of the time.” Cronbach's alpha for the MHI-3 in the INCPAD sample was 0.77.

Analysis

Cross-lagged panel analysis, a structural equation modeling technique, was used to test our two hypotheses. A latent variable approach was used to account for measurement error. The model included baseline and 3-month data and incorporated a linear structural equation for a continuous dependent variable fatigue at 3 months. The equation included depression at baseline and 3 months and fatigue at baseline. The intervention group assignment was entered into the equation as a control variable.

Specification of the indicators for the measurement model was carefully considered to minimize the potential for measurement error. For depression, instrument sum scores were used as indicators, in keeping with recommendations to use intact scales with established norms and psychometric properties as indicators of a latent construct when feasible (Little, Cunningham, Shahar, & Widaman, 2002). Because the INCPAD study focused primarily on depression and pain, there was only a single scale assessing fatigue. Consequently, we used item-level indicators from the scale to model fatigue rather than scale totals. The use of
multiple (versus single) indicators results in more reliable and valid estimates of latent constructs in SEM, even if this requires the use of individual scale items (Kline, 2005). In addition, the use of the four indicators provided by the vitality scale met the recommendations regarding number of indicators-per-construct to support latent-variable SEM (Kline, 2005).

The structural model was specified according to convention in cross-lagged panel data (Finkel, 1995; Greenberg, 2008; Kessler & Greenberg, 1981; Shadish et al., 2002). The structural pathways between fatigue at baseline and 3 months and between depression at baseline and 3 months served to adjust each 3-month variable for its corresponding baseline level; therefore the 3-month variables represent residualized change in fatigue and depressive symptoms. A significant correlation for a pathway from baseline of one symptom to the other symptom at 3 months would support our hypotheses regarding directional effects.

The analysis was conducted with maximum likelihood estimation using LISREL 8.8 (Jöreskog & Sörbom, 2008). Fit indices were selected a priori to determine goodness-of-fit. We included indices of absolute fit, the chi-square statistic ($\chi^2$; a non-significant $\chi^2$ suggests good fit), and the Standardized Root Mean Squared Residual (SRMR; with 0 indicating perfect fit and values lower than 0.06 indicating good fit); a parsimonious fit index, the Root Mean Square Error of Approximation (RMSEA; with 0 indicating perfect fit and cutoff of < 0.05); and an incremental fit index, the Comparative Fit Index (CFI; with values closer to 1 indicating better fitting model; cutoff > 0.95) (Hu & Bentler, 1999; Weston, Gore, Chan, & Catalano, 2008). The subject-to-parameter ratio was 9:1, providing adequate power for a viable analysis (Kline, 2005). Residuals between the same indicators across time-points (e.g., PHQ-9 at baseline and 3 months) were allowed to correlate freely. Alpha was set at .05, two-tailed.

Results

The sample size was reduced from 405 to 329 for the main analysis because we used only cases with full data. Cases that were not assessed at 3 months (17% of the sample) accounted for a large proportion of the missing data. At the time their 3-month assessment was due, 16 participants had died, 24 had dropped out, and the research team was unable to make contact with 30. Baseline characteristics of the 329 participants who were included in the current analysis were compared to those dropped from the analysis, and the results are presented in Table 1. No significant differences were found between the groups for age, gender, race, education, marital or employment status, income, number of comorbid diseases, disability, or mean scores for fatigue, pain, depression, or quality of life. The samples differed in the proportion of participants with breast cancer (32% in the sample used in the analysis, 16% of the dropped cases; $p = .007$) and the proportion in recurrent or progressive stage of disease (17% in the sample used in the analysis, 34% of the dropped cases; $p = .002$). Differences in other types of cancer or phase of treatment were nonsignificant.
Bivariate correlations, means, and SDs for all latent variable indicators across both time-points are presented in Table 2; correlations between the variables and intervention arm are also included. All correlations between latent variable indicators were significant and in the expected direction. Visual inspection revealed that intercorrelations among the depression scales tended to be higher than those among the fatigue items. As for the cross-sectional correlation of fatigue with depression, the total for the SF vitality scale score at baseline was found to have relatively strong and significant correlations with all three depression scales. For the PHQ-9, $r = -0.49$, for the SCL-20, $r = -0.62$, and for the MHI-3, $r = -0.54$. Negative correlations indicated that greater depression was associated with less vitality/greater fatigue.

Structural Analysis of Panel Data

A single analysis of panel data evaluated our hypotheses about associations between depression and fatigue in both directions. Several variables that have been found in previous studies to be associated with fatigue (Lawrence, Kupelnick, Miller, Devine, & Lau, 2004; Prue, Rankin, Allen, Gracey, & Cramp, 2006; Servaes, Verhagen, & Bleijenberg, 2002) were preliminarily evaluated as potential confounders. Variables included age, gender, pain, cancer type and phase, activity level, dyspnea, sleep quality, and anxiety. Preliminary tests were run to see if including each baseline variable in the model would improve the model fit indices or result in significant structural paths among the latent variables. No such effects were found; therefore, no control variables were included in the main analysis except the intervention group, which had been included in the hypothesized model.

The initial analysis with the a priori model specified as planned did not show adequate fit, $\chi^2 (76, N = 329) = 279.21, \text{ SRMR} = 0.073, \text{ RMSEA} = 0.09$, and $\text{CFI} = 0.97$. Examination of Lagrange multiplier modification indices, however, suggested that allowing the “full of life” fatigue indicator (sflife) to cross-load on depression at each time point would improve model fit. We considered whether this item (“How much of the time during the past 4 weeks have you felt full of life?”) could plausibly provide information relevant to both depression and fatigue and concluded that of the four items of the vitality scale, this one may have the greatest potential to provide information relevant to depressive symptoms. A new model was created with parameters freed from latent depression to the sflife indicator at each time point. Eight correlations between error terms were also freed based on modification indices. Upon finding that the first model didn't fit, we considered that the errors of similarly-worded items would likely be correlated. Parameters were freed to allow item error terms to correlate between “tired” and “wornout” items and between “full of life” and “a lot of energy” at each time point. Further, the “full of life” error terms were allowed to correlate with the MHI-3 error terms.

The revised model showed good fit, as demonstrated by three of the four fit indices: $\chi^2 (66, N = 329) = 88.16, p = 0.04, \text{ SRMR} = 0.030, \text{ RMSEA} = 0.032$, and $\text{CFI} = 1.00$. Although the $\chi^2$ was significant, this finding is expected since an exact fit is a rare occurrence, especially in larger samples (Weston et al., 2008). Kline (2005) advised that over-reliance on $\chi^2$ as a fit index may lead to rejection of models with reasonably good fit. Because the new measurement model fit the data well, it was used for the main analysis.
The cross-lagged structural path from baseline depression to fatigue at 3 months was nonsignificant ($\beta = 0.01, p = 0.92$) suggesting that baseline depression had no influence on change in fatigue after 3 months (see Figure 1). Likewise, the structural path from baseline fatigue to depression at 3 months was nonsignificant ($\beta = 0.10, p = 0.28$). A significant intervention effect was found for depressive symptoms, as those in the intervention arm showed greater 3-month decreases in depressive symptoms than those in the control arm ($\beta = 0.20, p < 0.0001$). No effect of the intervention was found for fatigue, however, as the structural path from the intervention control variable to fatigue was nonsignificant ($\beta = -0.09, p = 0.11$).

An alternative model was evaluated to explore the consequences of reducing obvious measurement overlap between fatigue and depression. In this analysis, two items of the SCL-20 and one from the PHQ-9 that assess fatigue/energy were dropped. Model indices suggested a good fit, and unlike the main model, the $\chi^2$ was nonsignificant, suggesting it met the specifications for absolute fit, $\chi^2 (66, N = 329) = 81.67, p = 0.09$, SRMR = 0.027, RMSEA = 0.044, and CFI = 1.00. The cross-lagged paths between fatigue and depression remained nonsignificant, suggesting no directional relationship—consistent with the main analysis. The structural effects did not change in any apparently meaningful way. This outcome suggests the inclusion of fatigue items on the depression scales did not unduly influence the findings.

Discussion

No evidence was found to support a hypothesis of a directional relationship between depression and fatigue in the cancer context. Our structural equation model fit the data reasonably well. Depression and fatigue were strongly correlated in the sample, and depression and fatigue decreased on average over the course of the months, yet cross-lagged relationships were not detected.

These findings contribute to what is currently understood about the interrelationship of depression and CRF. Although a strong correlation between depression and CRF has been established in previous research, few longitudinal studies have been undertaken and little is known about directional relationships. Only a few studies have attempted to evaluate whether these symptoms tend to change together or independently over time, and the results have been mixed and inconclusive. To our knowledge, this is the first study to examine the bidirectional relationships among depressive symptoms and fatigue using structural equation modeling.

The current findings are consistent with the majority of the published longitudinal studies in this area. The study with the most notable similarities in findings is a randomized, double-blind controlled trial of paroxetine to treat fatigue in patients ($N = 479$) undergoing chemotherapy for the first time (Morrow et al., 2003). Fatigue was the targeted symptom in this trial, whereas the INCPAD trial targeted depression and pain. The study by Morrow and colleagues tested whether fatigue and depression share a common etiology (i.e., a neural pathway involving serotonin). However, they found that paroxetine, a selective serotonin reuptake inhibitor, reduced symptoms of depression within 3 weeks but had no effect on
fatigue at either of two follow-up assessments at cycles 3 and 4 of chemotherapy. As in the current study, these investigators found that, although depression and fatigue were strongly correlated in the sample, only depression was affected by the intervention. In five other longitudinal studies, researchers reported findings that suggest that the fatigue and depressive symptoms change (or fail to change) independently over time (Geinitz et al., 2001; Pirl, 2008; Schumacher et al., 2002; Stone et al., 2000; Visser & Smets, 1998).

Tchekmedyian and colleagues (2003) (N = 250) reported finding a longitudinal relationship in change in depression and fatigue, which could be considered to contradict the current findings. Their study, a secondary analysis of data from a randomized, double-blind, placebo-controlled trial of darbepoetin alfa to treat anemia in lung cancer patients, was conceived to test of whether improvements in fatigue would predict improvements in the distress of depression and anxiety. Assessment of fatigue, anxiety, and depression occurred at baseline and between weeks 4 and 12. Improvements in fatigue were associated with reductions in both depression and anxiety. In multiple regression, for each unit of improvement in the fatigue score, there was a corresponding improvement of 0.7 units in anxiety and 0.8 units in depression levels. Although the Tchekmedyian and colleagues study findings suggest a directional influence between fatigue and depression, it is possible that the fatigue associated with anemia is different from non-anemic CRF. For example, anemia has been established as a “treatable contributing factor” of CRF and has a pharmacologic treatment that is known to be helpful (Berger et al., 2010). Unlike the Tchekmedyian et al. study, presence of anemia was not an inclusion criterion in the current study and it was neither assessed nor treated. These differences may account for the variation in findings between the current study and the previous study. Overall, in terms of the few relevant longitudinal studies that have been found in the literature, none directly contradicted the current study.

Implications

Theory—The current study adds new data to the evidence regarding the three unanswered questions about CRF that have been posed by Jacobsen and colleagues (2003): first, to what degree do fatigue and depression conceptually differ; second, to what degree do they co-occur; and, third, are there causal relationships? Regarding the first question, our findings are consistent with the theory that depression and fatigue are two distinct entities in the cancer context. The differential response to the intervention that specifically targeted depression and pain but not fatigue stands as a compelling demonstration of two distinct symptoms. Intervention arm membership had a significant effect on change in 3-month depression while it had no demonstrated effect on change in fatigue. Moreover, depression changed more than fatigue over the 3 months.

As for the second question regarding co-occurrence of CRF and depression, the current findings lend further support to the well-established recognition of a strong association. The correlations in our relatively large sample of patients with various types of cancer were strong at baseline. Although the magnitude of the association between fatigue and depression declined over time as depression improved and fatigue did not, the correlation remained strong.

Health Psychol. Author manuscript; available in PMC 2014 December 01.
Regarding the third question, we found no support for a lagged relationship between the two symptoms in either direction. What, then, might be the explanation for the strong correlation? As was suggested by Jacobsen and colleagues (2003) (Jacobsen et al., 2003), a third factor or set of factors may cause both fatigue and depression in the cancer context. Numerous possible common causes have been advanced speculatively, such as certain forms of cancer treatment (including biological response modifiers), increased levels of proinflammatory cytokines that occur as a result of cancer treatment, and certain types of cancer such as pancreatic (Bower, 2007; Fann et al., 2008; Jacobsen et al., 2003). These potential third factors could account for the strong cross-sectional associations between depression and fatigue, such as those we observed in the INCPAD sample.

**Practice**—Our findings are consistent with suggestions in the literature that treating depression may not be sufficient in ameliorating fatigue. NCCN treatment guidelines for CRF state that antidepressants are not recommended (Berger et al., 2010). Although at least two other reviews of treatments for CRF endorsed antidepressants, neither review provided empirical support for this recommendation. One cited only “clinical observation” and no clinical trials to support the recommendation (Jacobsen et al., 2003); the other cited three placebo-controlled randomized trials of antidepressants that failed to improve CRF, but the authors nevertheless recommended antidepressants as potentially helpful for fatigue if it is comorbid with depression (Escalante, 2010). The current study lends evidence to support NCCN’s current guidelines.

The findings are supported by several strengths in study design. The majority of the sample entered the study with both elevated depression and fatigue, assuring enough variation in each variable to demonstrate change over time. The findings were consistent with those of the INCPAD trial, suggesting that the depression-and-pain intervention had positive effects for depression compared to the usual care control group. Fatigue, however, was not significantly affected by intervention group membership, according to the panel analysis. The sample of 329 was sufficient to support the analysis. Moreover, because latent variable SEM accounts for measurement error, the validity of these findings may be less threatened by measurement error than are those of studies using other statistical approaches.

**Limitations**

The study sample included only cancer patients that met criteria for at least moderate levels of depression and/or pain; thus, our results may not extend to cancer patients who are not experiencing clinically significant levels of these symptoms. However, systematic reviews have shown that pain is present in more than half of patients with cancer, and depression is present in at least a quarter to a third of patients (Bottomley, 1998; Carr et al., 2002; Kim, Dodd, Aouizer, Jahan, & Miaskowski, 2009; van den Beuken-van Everdingen et al., 2007). Nonetheless, replicating our findings in a sample that also included more patients without pain or depression would broaden the generalizability of our path analysis. The use of a heterogeneous sample of cancer patients—with a range of types of cancer and in various phases of treatment or survivorship—can be considered both a strength and a limitation of our study. On the one hand, our sample reflects the wide range of patients treated in oncology clinics, which enhances generalizability to a clinical setting. On the other hand, the
etiology of CRF and its relationship to depression may vary across types of cancer and phases of treatment. In the current study, mean scores for the vitality scale did not differ by cancer type or treatment phase. Our sample size, however, was insufficient to evaluate differential effects by cancer type or treatment phase. Therefore, our findings may not apply to all types of cancer or phases of treatment.

Another limitation is the lack of multiple well-validated fatigue scales to support the latent variable fatigue in the current analysis. Having three such scales for depression was a strength. The lack of multiple full scales for fatigue was appropriately addressed by validating the vitality scale in comparison with the Fatigue Symptom Inventory in a subsample. That plus the acceptable performance of the fatigue indicators in the CFA phase of the analysis suggests measurement was at least adequate in this analysis.

A final issue to take into account when considering our findings is the length of the follow-up period. Although three months represents the acute phase of depression treatment and is the period during which the greatest degree of response to antidepressant therapy typically occurs, the current study cannot evaluate whether changes in depression predict changes in fatigue at time points shorter or longer than three months. It is possible that we would have detected lagged relationships between fatigue and depressive symptoms over a shorter period, as there is significant variability in CRF between and within days (Jacobsen, 2004). Moreover, circulating proinflammatory cytokines, which may give rise to both fatigue and depressive symptoms (Bower, 2007), also tend to be highly variable within days (Miller, Ancoli-Israel, Bower, Capuron, & Irwin, 2008). Similarly, our findings might be different for follow-up periods longer than 3 months.

**Future Directions**

A natural next step that would add confidence to the findings regarding directional effects is to extend the current analysis to a third wave of panel data. Three-wave and multiwave panels have been suggested as optimal for causal analysis (Finkel, 1995). Ideally, the current findings should be replicated in larger samples and with multiple validated scales for fatigue. A larger heterogeneous sample similar to that of the INCPAD study would provide power to detect smaller effects. More importantly, a much larger sample would allow the model to be tested by type and phase of cancer, thereby helping to resolve the unanswered question as to whether depression may have a directional influence on CRF in specific types or phases of cancer. A larger sample could more readily be extended to multiple waves. It would also be informative to complete a similar study with a nonpharmacologic intervention for fatigue and test for a lagged relationship with depression.

The current study underscores the need for future research focused on understanding CRF. Given its prevalence and the associated disability, surprisingly little is known about the etiology of CRF and how to help patients who suffer from it. The current findings add to literature on this under-studied but important condition, and it is hoped that our study will help to inform the development of effective treatments for CRF.
Acknowledgments

This manuscript was supported by a grant from the National Cancer Institute to Dr. Kroenke (R01 CA-115369) and a grant from the National Cancer Institute (R25 CA-117865-01A11)

References


Cuijpers P, Smits N, Donker T, ten Have M, de Graeff R. Screening for mood and anxiety disorders with the five-item, the three-item, and the two-item Mental Health Inventory. Psychiatry Research. 2009; 168(3):250–255. [PubMed: 19183534]


Escalante CP. Cancer-related fatigue: Treatment. 2010; 18 UpToDate.


Health Psychol. Author manuscript; available in PMC 2014 December 01.


Health Psychol. Author manuscript; available in PMC 2014 December 01.


Figure 1.
Model of the longitudinal relationships between fatigue and depression in cancer patients with depression and/or pain. *Shaded indicator (life) is a crossloading item. Values for unidirectional arrows (structural paths) are standardized regression coefficients; values associated with bidirectional arrows are Pearson correlation coefficients. Paths with significant coefficients are solid; nonsignificant paths are dashed. The intervention group was entered as a control variable using dummy coding (intervention group = 1; control group = 2). Loadings of fatigue indicators are negative because higher observed scores suggest lower fatigue (tired and worn scores were reversed). p < 0.05.
<table>
<thead>
<tr>
<th>Baseline Characteristics *Bold value in range indicates worse score</th>
<th>Panel Analysis Sample n = 329</th>
<th>Participants w/Missing Data n = 76</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD) age, yr</td>
<td>58.5 (10.5)</td>
<td>60.1 (12.0)</td>
<td>.10</td>
</tr>
<tr>
<td>Female sex, n (%)</td>
<td>228 (69)</td>
<td>47 (62)</td>
<td>.26</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>260 (79)</td>
<td>62 (82)</td>
<td>.76</td>
</tr>
<tr>
<td>Black</td>
<td>61 (18)</td>
<td>12 (16)</td>
<td>.69</td>
</tr>
<tr>
<td>Other</td>
<td>8 (2)</td>
<td>2 (3)</td>
<td>.76</td>
</tr>
<tr>
<td>Education, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than High school</td>
<td>68 (21)</td>
<td>19 (25)</td>
<td>.50</td>
</tr>
<tr>
<td>High school</td>
<td>131 (40)</td>
<td>29 (38)</td>
<td>.89</td>
</tr>
<tr>
<td>Some college or trade school</td>
<td>89 (27)</td>
<td>19 (25)</td>
<td>.83</td>
</tr>
<tr>
<td>College graduate</td>
<td>41 (13)</td>
<td>9 (12)</td>
<td>.96</td>
</tr>
<tr>
<td>Married, n (%)</td>
<td>159 (48)</td>
<td>33 (43)</td>
<td>.52</td>
</tr>
<tr>
<td>Employment status, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>65 (20)</td>
<td>17 (23)</td>
<td>.72</td>
</tr>
<tr>
<td>Unable to work due to poor health or disability</td>
<td>148 (45)</td>
<td>28 (37)</td>
<td>.25</td>
</tr>
<tr>
<td>Retired</td>
<td>92 (28)</td>
<td>25 (33)</td>
<td>.48</td>
</tr>
<tr>
<td>Other</td>
<td>24 (7)</td>
<td>6 (8)</td>
<td>.95</td>
</tr>
<tr>
<td>Mean (SD) scores *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF Vitality Scale Total (score range, 0-100)</td>
<td>29.12 (19.6)</td>
<td>24.51 (17.0)</td>
<td>.25</td>
</tr>
<tr>
<td>BPI pain severity (score range, 0-10)</td>
<td>4.20 (2.3)</td>
<td>4.54 (2.5)</td>
<td>.23</td>
</tr>
<tr>
<td>SCL-20 depression (score range, 0-4)</td>
<td>1.43 (0.7)</td>
<td>1.50 (0.6)</td>
<td>.42</td>
</tr>
<tr>
<td>Type of cancer, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>106 (32)</td>
<td>12 (16)</td>
<td>.007</td>
</tr>
<tr>
<td>Lung</td>
<td>60 (18)</td>
<td>21 (28)</td>
<td>.09</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>57 (17)</td>
<td>13 (17)</td>
<td>.90</td>
</tr>
<tr>
<td>Lymphoma and hematological</td>
<td>41 (13)</td>
<td>12 (16)</td>
<td>.56</td>
</tr>
<tr>
<td>Baseline Characteristics *Bold value in range indicates worse score</td>
<td>Panel Analysis Sample n = 329</td>
<td>Participants w/Missing Data n = 76</td>
<td>P Value</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Genitourinary</td>
<td>33 (10)</td>
<td>8 (11)</td>
<td>.94</td>
</tr>
<tr>
<td>Other</td>
<td>32 (10)</td>
<td>10 (13)</td>
<td>.50</td>
</tr>
<tr>
<td>Phase of cancer, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newly-diagnosed</td>
<td>126 (38)</td>
<td>24 (31)</td>
<td>.37</td>
</tr>
<tr>
<td>Maintenance or disease-free</td>
<td>146 (44)</td>
<td>26 (34)</td>
<td>.14</td>
</tr>
<tr>
<td>Recurrent or progressive</td>
<td>57 (17)</td>
<td>26 (34)</td>
<td>.002</td>
</tr>
</tbody>
</table>
Bivariate Correlations, Means, and Standard Deviations of Latent Variable Indicators Across Both Time-Points

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. sfenerT1</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. sflifeT1</td>
<td>.41</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. sfwornT1</td>
<td>.48</td>
<td>.29</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. sftireT1</td>
<td>.49</td>
<td>.35</td>
<td>.74</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. PHQ9T1</td>
<td>-.36</td>
<td>-.39</td>
<td>-.34</td>
<td>-.40</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. SCL20T1</td>
<td>-.47</td>
<td>-.46</td>
<td>-.46</td>
<td>-.51</td>
<td>.74</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. MHI-3T1</td>
<td>-.37</td>
<td>-.53</td>
<td>-.33</td>
<td>-.40</td>
<td>.65</td>
<td>.74</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. sfenerT2</td>
<td>.41</td>
<td>.29</td>
<td>.39</td>
<td>.36</td>
<td>-.27</td>
<td>-.35</td>
<td>-.27</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. sflifeT2</td>
<td>.39</td>
<td>.41</td>
<td>.26</td>
<td>.30</td>
<td>-.34</td>
<td>-.42</td>
<td>-.43</td>
<td>.53</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. sfwornT2</td>
<td>.33</td>
<td>.30</td>
<td>.44</td>
<td>.43</td>
<td>-.36</td>
<td>-.41</td>
<td>-.33</td>
<td>.53</td>
<td>.49</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. sftireT2</td>
<td>.32</td>
<td>.25</td>
<td>.38</td>
<td>.43</td>
<td>-.24</td>
<td>-.31</td>
<td>-.25</td>
<td>.56</td>
<td>.45</td>
<td>.78</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. PHQ9T2</td>
<td>-.31</td>
<td>-.25</td>
<td>-.31</td>
<td>-.34</td>
<td>.44</td>
<td>.53</td>
<td>.46</td>
<td>-.48</td>
<td>-.54</td>
<td>-.60</td>
<td>-.54</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. SCL20T2</td>
<td>-.32</td>
<td>-.26</td>
<td>-.30</td>
<td>-.28</td>
<td>.43</td>
<td>.58</td>
<td>.43</td>
<td>-.51</td>
<td>-.59</td>
<td>-.58</td>
<td>-.50</td>
<td>.85</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. MHI-3T2</td>
<td>-.33</td>
<td>-.30</td>
<td>-.27</td>
<td>-.30</td>
<td>.44</td>
<td>.53</td>
<td>.56</td>
<td>-.44</td>
<td>-.63</td>
<td>-.51</td>
<td>-.46</td>
<td>.73</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Tx Group</td>
<td>-.02</td>
<td>.03</td>
<td>.07</td>
<td>.06</td>
<td>.04</td>
<td>-.01</td>
<td>.04</td>
<td>-.05</td>
<td>-.11</td>
<td>-.03</td>
<td>-.06</td>
<td>.16</td>
<td>.20</td>
<td>.12</td>
<td>—</td>
</tr>
</tbody>
</table>

M     | 1.90| 2.40| 2.29| 2.06| 12.84| 1.44| 8.05| 2.19| 2.58| 2.65| 2.44| 9.20| 1.10| 7.26 |
SD    | 0.88| 1.14| 1.06| 0.99| 6.85| 0.73| 2.96| 1.03| 1.22| 1.13| 1.08| 6.28| 0.69| 3.02 |

Note. Fatigue indicators are sfener ("lot of energy"), sflife ("full of life"), sfworn ("feel worn out"), sftire ("feel tired"). Depression indicators are PHQ9 (Patient Health Questionnaire, Depression), SCL20 (Depression Subscale of the Hopkins Symptom Checklist), MHI-3 (Mental Health Inventory Depression Subscale). T1 = baseline; T2 = 3 months. Negative correlations suggest a positive association between fatigue and depression because higher vitality scores suggest less fatigue whereas higher depression scores suggest worse symptoms. Correlations between variables 1 – 14 are all significant at the 0.01 level (2-tailed). For variable 15, all correlations are nonsignificant except for 9, 12, 13, and 14.