Evaluating the impact of depression screening in general hospitals: a pilot study of patient reported outcomes of quality of life, functioning, and psychiatric symptom severity

Waguih William IsHak [1, 2] *****, **Enrique Lopez** [1, 2], **Alexander Joseph Steiner** [1], **Payam Kharazi** [1], **Paula Cedillo** [1],
Marcia Nimmer [1], **Nardos Bellete** [1], **Ariella Benji** [1], **Robert Berberyan** [1], **Kimberly Smith** [1], **Lancer Naghdechi** [1],
Karim Yahia Jaffer, **Linda Parisi** [1], **Linda Burnes-Bolton** [1], **Itai Danovitch** [1, 2]

[1] *Department of Psychiatry and Behavioral Neurosciences, Cedars-Sinai Medical Center, Los Angeles, California*
[2] *Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine,*
University of California Los Angeles, Los Angeles, California
*Corresponding author E-mail: Waguih.IsHak@cshs.org

Abstract

Background: Depression is common among patients with medical co-morbidities and has an adverse impact on wide ranging health outcomes. Depression is particularly prevalent in general hospital settings, where it has been linked to worse course of illness and higher rates of hospital readmission. While depression screening has been established as a best practice in outpatient settings, far less is known about the utility of screening for depression in inpatient medical settings.

Aims: To determine whether screening for depression in a tertiary care medical hospital produces reliable and valid results. To identify obstacles and success factors that impact utility of depression screening. To evaluate the relationship between depression screening scores and patient-centered outcomes such as: functioning and quality of life.

Method: Participants were 30 patients who were admitted to a tertiary care medical hospital, had a depression screening assessment completed by nursing on admission, and for whom a psychiatry consultation was requested. Patient responses to several measures were recorded and compared to PHQ screening on admission.

Results: Depressed patients had significantly worse scores compared to non-depressed patients on the PHQ, WHO-5, PROMIS scores (global health, global mental health, physical functioning, anxiety, and fatigue), and Charleston Comorbidity Index (all p values < .05). Correlational analyses provided the strength of the relationships for each measures test-retest reliability.

Conclusions: Findings support the clinical utility of screening for depression during inpatient hospital admissions. Assessing QOL, functioning, and psychiatric symptoms can help identify risk factors negatively influencing treatment outcomes.

Keywords: *Quality of Life; Functioning; Hospital Screening; Data Registry; Patient-Reported Outcomes.*

1. Introduction

Psychiatric symptoms/disorders affect nearly one in five adults in the United States. Specifically, 43.7 million or 18.6% of adults experience mental illness in a given year (NIMH 2016). According to the Center for Disease Control, mental illness is associated with higher rates of chronic diseases, including diabetes, asthma, obesity, cancer and cardiovascular disease (CDC 2016). The scientific literature base clearly outlines that psychiatric symptoms and disorders are strongly associated with impaired health outcomes in medical and surgical patients including length of stay, readmission rates, direct and indirect costs of healthcare, patient satisfaction, social/occupational functioning, quality of life (QOL), and overall health outcomes (Bonicatto et al. 2001, Chapman et al. 2005, Evan et al. 2005, Kessler et al. 2008, El-Gabalawy et al. 2010).

Additionally, research suggests that depression significantly impacts hospitalized patients on multiple levels (Verbosky et al. 1993). Specifically, depressed patients experience longer recovery times and are more likely to be readmitted within 30 days after being discharged. This often increases patients' depressive symptoms and lowers their overall functioning and QOL. Subsequently, this adds more financial strain to the patient and the health system, as medical insurance does not always cover readmission cycles (Katon 1996). These findings reinforce the need to integrate the traditional medical model, which emphasizes the recognition and reduction of symptoms, with a more comprehensive definition of health that emphasizes physical, mental, and social well-being (Armitage et al. 2009, Enthoven 2009).

Recognizing that psychiatric factors play into a patient's health and recovery process, Cedar-Sinai Medical Center has begun screening patients for depression. With this screener, the health care team is able to detect and address any depressive symptoms that may hinder a patient's recovery, extend their length of stay, or contribute to their readmission within a short window of time. Upon admission to the hospital, all patients are screened for depression by nursing staff using the 2-item Patient Health Questionnaire (PHQ-2) (Kroenke et al. 2003). Patients who screen positive for depression are brought to the attention of the admitting physician in addition to activating a social work consult, often

leading to a psychiatric consultation and treatment initiation using biopsychosocial interventions.

The different factors that could potentially affect patient health and well-being were recorded and monitored as part of a data registry. Patients in this sample were admitted for medical reasons at Cedar-Sinai Medical Center and were given self-report health measures to assess QOL, functioning, and psychiatric symptom severity as a part of routine clinical care from May 2014 through May 2015. Patients were asked to report on their functioning, satisfaction with their social role, sleep, pain interference, anxiety, and depression through the Patient Reported Outcome Measurement Information System (PROMIS-29) (Cella et al. 2010) and the World Health Organization Disability Assessment Schedule (WHODAS 2.0) (Üstün et al. 2010). Their responses were recorded and coded, to see how effectively the measures could capture the factors affecting overall QOL, and how well the different measures correlated with each other as well as with the nurses' intake evaluations.

The aim of this study is to identify the factors that impede patient recovery and impact their QOL and function, as well as identify the best measures for assessing these factors. An additional goal of the current study is to ensure the measures are used to effectively identify patients experiencing depression and other mental health challenges, and to promptly intervene and prevent a patient's mental health from hindering the recovery of their physical health and negatively impacting their overall QOL.

2. Methods

2.1. Participants

Potential subjects were identified by their treating physicians or by positive endorsement of depressive symptoms on the 2-item Patient Health Questionnaire (PHQ-2) (Kroenke et al. 2003) depression risk screening measure given by nursing staff upon admission and referred to researchers. No patients were excluded for demographic background, language, or education level. While most patients presenting at Cedars-Sinai Medical Center reside in the Greater Los Angeles region, a substantial portion of admitted patients are either from other regions close to the Los Angeles area, from across the nation, or live internationally and are seeking specialty medical treatment. The lead physician investigator and neuropsychologist determined a patient's capacity to consent for participation in the study. Exclusion criteria included patients with less than 18 years of age at the time of admission, patients under conservatorship, patients suffering from dementia or florid psychosis, or patients otherwise lacking capacity to make medical decisions, as these individuals are considered unable to fully understand and consent to research.

2.2. Study procedures and data collection

The Cedars-Sinai Medical Center Institutional Review Board (IRB) in Los Angeles, California, USA, approved this study. All research conducted followed the guidelines established in the Declaration of Helsinki for ethical research with human participants. The preliminary data referenced in this article includes participants who signed informed consent during the time period of May 2014 to May 2015. Patients who met all inclusion and exclusion criteria were asked to participate in the research study by their psychiatrist. Patients were informed that their participation was voluntary and that they were free to withdraw their consent at any time during or after the evaluation process, and that their decision had no effect on their treatment or relationship to Cedars-Sinai Medical Center. Data on QOL, overall health, demographic and clinical data were collected, de-identified, and stored in the registry for research analysis. In order to maintain the confidentiality of subjects all data collected was securely maintained and only accessible to certified-study staff listed in the Cedars-Sinai IRB approved application.

2.3. Clinical measures

2.3.1. Patient health questionnaire (PHQ)

The PHQ-9 (Kroenke et al. 2001) measures the nine depressive symptoms delineated in the Diagnostic and Statistical Manual (DSM-IV) (APA 2000). Each question is rated on a 4-point Likert scale ranging from 0 "not at all" to 3 "nearly every day"; the total score is the sum of all nine items. Scores may fall into the following categories: 1-4 (minimal depression), 5-9 (mild depression), 10-14 (moderate depression), 15-19 (moderately severe depression), or 20-27 (severe depression). Patients who scored > 12 on the PHQ were deemed "depression screen positive". The PHQ-9 has an excellent internal consistency reliability with a Cronbach's α range varying from 0.86 - 0.89, and a test-retest reliability of 0.84 (Kroenke et al. 2001).

2.3.2. PROMIS - global health scale (GHS)

The NIH PROMIS GHS measures patient–reported health and QOL by assessing five primary domains: physical function, fatigue, pain, emotional distress, and social health (Cella et al. 2010). The GHS consist of two factors, a global physical health factor and a global mental health factor. Items employ a 5-point Likert scale. The GHS takes approximately 5 minutes for participants to complete and consist of 10 items. Raw scores are converted to standardized t-scores, each with a mean of 50 (SD=10) where higher or lower scores indicate better or worse health than the general population. While the GHS has internal consistency reliability of 0.92, one study conducted a factor analysis and found the single-factor structure was rejected statistically and did not fit the data (Hays et al. 2009). Therefore, we looked at the MHS and PHS scores separately.

2.3.3. PROMIS – global mental health score (MHS)

The PROMIS MHS score is derived from 4 items on the GHS related to mental health, QOL, satisfaction with social activities, and emotional problems with a 5-point Likert scale for each item (Cella D et al. 2010). Raw scores are converted to standardized t-scores, each with a mean of 50 (SD=10) where higher or lower scores indicate better or worse health than the general population. The MHS has good internal consistency reliability with a Cronbach's α of 0.86 (Hays et al. 2009).

2.3.4. PROMIS- global physical health score (PHS)

The PROMIS PHS score is also derived from the GHS and consist of 4 items of overall physical health, pain, fatigue, and functioning with each a 5-point Likert scale for each item (Cella et al. 2010). Raw scores are converted to standardized t-scores, each with a mean of 50 (SD=10) where higher or lower scores indicate better or worse health than the general population. The PHS has good internal consistency reliability with a Cronbach's α of 0.81 (Hays et al. 2009).

2.3.5. PROMIS-29

Funded by the National Institutes of Health (NIH), the PROMIS measures were developed using modern psychometric techniques (Cella et al. 2010). Specifically, item banks were established using item-response theory to generate a comprehensive set of calibrated items for all of the domains, and have subsequently been found to retain strong psychometric properties (high reliability and validity), with large samples reflecting 2000 census demographics (Cella et al. 2010). The majority of items (28/29) employ a 5-point Likert scale response (e.g., 1 = not at all, 2 = a little bit, 3 = somewhat, 4 = quite a bit, 5 = very much). In order to standardize raw scores, raw scores are summed and then converted to t-scores for each domain, where a score of 50 represents the average population mean value (50th percentile), and one standard deviation (SD) equals 10 points. The PROMIS-29 measures consist of sev-

eral domains including: physical function (4 items), anxiety (4 items), depression (4 items), fatigue (4 items), sleep disturbance (4 items), satisfaction with social roles (4 items), pain interference (4 items), and one pain intensity item scored on a 10-point range, where the raw score is used for qualitative analysis. This study utilized the PROMIS-29 short forms and scored items based on the hand-scoring guidelines found on the PROMIS website (nihpromis.org).

2.3.6. PROMIS-cognitive abilities scale

Like the PROMIS-29, the 4-item PROMIS Applied Cognitive-Abilities scale was developed using item-response theory as a brief but valid and reliable tool that is standardized and assesses subjective ratings of cognitive functioning during the prior 7 days. Participants rate their responses on a 5-point Likert scale ranging from 1 (not at all) to 5 (very much). A total raw score is calculated by the sum of all the participants' responses (range = 0 to 40), and then converted to a t-scores, where a score of 50 represents the average population mean value (50th percentile), and one standard deviation (SD) equals 10 points. Prior research of the 4-item PROMIS-Cognitive Abilities scale with medical populations has found remarkable internal consistency reliability with a Cronbach's α of 0.95 (Saffer et al. 2015).

2.4. CAGE-AID

The CAGE-AID is a 4-item substance abuse screening measure adapted from the original CAGE to include drug use as well as alcohol use. First published by Ewing, the acronym was made from its four questions: Cut-Down, Annoyed, Guilty, and Eye-Opener (Ewing 1984). One diagnostic meta-analysis found that for a cutoff point > 2, the pooled sensitivity of the measure was far better for inpatients (0.87) populations, compared to primary care patients (0.71), or ambulatory patients (0.60), and that the pooled sensitivity varies by diagnostic group, with a pooled area under the curve (AUC) of 0.87 (Aertgeerts et al. 2004). A more recent review study of the CAGE questionnaire found the CAGE had high test-retest reliability (0.80 - 0.95) (Dhalla et al. 2007).

2.5. Psychiatric consultation identification (PCI)

The PCI score was developed at Cedars-Sinai Medical Center (CSMC) to systematically and proactively identify medical-surgical inpatients with significant underlying mental conditions. The score is based on 40 risk factors from APM practice guidelines and is used to screen hospitalized patients with medical conditions who might benefit from a psychiatric consultation. With this instrument Cedars-Sinai strives to improve the overall detection of psychiatric comorbidity that interferes with effective medical/surgical management (IsHak et al. 2014).

2.6. Charleston comorbidity index (CCI)

The CCI is a method of categorizing comorbidities of patients based on the International Classification of Diseases (ICD) diagnosis codes (Charleson et al. 1987). A weighted index accounts for the number and seriousness of various comorbid diseases, while the index has been found to be a reliable predictor of death after 10 years. Each comorbidity category has an associated weight based on adjusted risk of mortality and with each increased level on the CCI, there are stepwise increases in the cumulative mortality attributed to comorbid illness and scores are calculated as the sum of all the weights, resulting in a single comorbidity score for a patient (Charleson et al. 1987).

2.7. World health organization disability assessment schedule (WHODAS) 2.0

The WHODAS 2.0 and its brief version 12-item WHODAS 2.0 measure functioning in: Cognition –understanding and communi-

cating; Mobility– moving and getting around; Self-care– hygiene, dressing, eating and staying alone; Interactions with others– interacting with other people; Life activities– domestic responsibilities, leisure, work and school; and Participation– joining in community activities (Üstün et al. 2010). Scoring utilizes one of two methods: simple scoring involves the sum of all score, and complex scoring uses a script converting the score using item-response theory to a range from 0 (no disability) to 100 (total disability). The WHODAS 2.0's administration time is approximately 5 minutes and it has a stable factor structure, high test-retest reliability (intraclass correlation coefficient = 0.98), and good internal consistency reliability with a Cronbach's α of 0.86 (Üstün et al. 2010).

2.8. World Health Organization-5

The World Health Organization (WHO) Regional Office in Europe presented this measure in 1998 as results from an on-going well-being study done in multiple member countries worldwide. Originally designed as a 28-item measure, then scaled down to a 10-item measure, the current 5-item questionnaire measures positive mental health and includes the following five content areas: cheerfulness, calm, vigor, rest, and interest in daily pursuits (WHO 1998, Bech 2012, Bech et al. 2013). The 5-item questionnaire measures well-being in the last two weeks and is rated along a 6-point Likert scale with scores ranging between 0 "at no time" to 5 "all of the time." The raw score is calculated by summing the scores for all 5 items (possible range = 0-25), with 0 representing the most impaired QOL and 25 representing the best possible QOL score. A percentage score can be obtained by multiplying the total raw score by 4, where again lower scores indicate worse QOL (Bech 2004). The internal validity of the WHO-5 using non-parametric Mokken analysis has been reported to have a Loevinger coefficient of homogeneity of 0.5 or greater, which indicates strong support of scalability (Bech 2004). The WHO-5 has been found to be a preferred screening tool for depression, with one study reporting good sensitivity (.93), and negative predictive value (.98) but relatively weaker specificity (.64) and positive predictive value (.34), although it is important to note that screening measures are developed to identify at risk groups for a given disorder and the most important operating characteristics are negative predictive value and sensitivity (Henkel et al. 2003).

2.9. Statistical analysis

Analyses were conducted to assess for violations of normality of distribution using a Levene's test of equal variances. Summary values are expressed as means (M) and standard deviations (SD) for continuous variables and frequencies (%) for categorical variables. We considered tests with p values < .05 to be statistically significant. We examined the differences in the clinical measures across 2 groups using an analysis of variance (ANOVA) for multiple comparisons: (1) PHQ Positively and Negatively screened for depression; (2) Depressed versus Non-Depressed groups utilizing a t-score cutoff > 65 on the PROMIS-Depression for Depressed group. Mann-Whitney test were used for group comparisons if normality distribution assumptions were violated. We performed Chi-square analysis on PHQ groups (screened positively or negatively by Nurse) with Depressed Group versus Non Depressed group based on a t-score cutoff > 65 on the PROMIS-Depression for Depressed group. We used Pearson's correlation coefficients (r) to determine the strength of the relationships between total scores of the measures on both the PHQ screened by Nurse and the PHQ retest data. Analyses were performed using SAS software version 9.2 (SAS Institute Inc., Cary, NC, USA).

3. Results

The demographic and clinical characteristics of the study population are presented in Table 1. Among the total number of participants included in the study (n = 30), there were equal number of

women and men, 66.7% were Caucasian, and 60% were single, with a mean age of 42.37 (SD = 14.85). Approximately one third to one fourth of the sample subjects were either employed (33.3%) or disabled (26.7%) respectively. About half of the sample scored positive on the nurse administered PHQ-9 depression inventory (53.3%). All raw scores on psychiatric measures for entire population are displayed in Table 2.

Table 1: Sample Demographics: Age, Gender, Marital Status, Ethnicity, Employment Status, PHQ Screen (N=30)

Variable	M	SD
Age	42.37	14.85
	%	
Female	50.0	
Married	16.7	
Single	60.0	
Divorced	13.3	
Unknown/Other	11.0	
Caucasian	66.7	
African American	23.3	
Asian	6.7	
Other	4.3	
Employed	33.3	
Unemployed	23.3	
Disabled	26.7	
Student	3.3	
Retired	3.3	

PHQ Nurse Screen Positive 53.3 (16/28), 2 individuals were not screened

Table 2: Mean Scores and Standard Deviations on All Subjects (N= 30) of the Psychiatry Measures

Variable	M	SD
PHQ Nurse (PHQ-N)	16.18	6.70
PHQ Retest	16.81	7.34
Well-Being Index (WHO-5)	8.04	6.12
PROMIS Global Health Score (GHS)	9.69	3.58
PROMIS Global Mental Health Score	8.17	3.46
PROMIS Physical Function	10.25	5.93
PROMIS Anxiety	14.21	4.04
PROMIS Depression	13.71	4.97
PROMIS Fatigue	15.18	4.88
PROMIS Sleep Disturbance	14.71	4.32
PROMIS Satisfaction with Social Role	7.79	4.64
PROMIS Pain Interference	15.21	5.44
PROMIS Pain Intensity	6.57	2.94
PROMIS Cognitive Abilities	9.85	5.07
GAGEAID Total	0.80	1.41
Psychiatry Consultation Index (PCI)	12.35	4.78
Charleston Comorbidity Index (CCI)	2.38	2.43

An ANOVA to compare clinical measures between patients who screened positive for depression (PHQ-9 score > 12) to those who screened negative demonstrated that the WHO-5, PROMIS Global Mental Health, PROMIS Anxiety, PROMIS Depression, PROMIS Fatigue, and PROMIS Cognitive Abilities measures were all significantly different between both groups (all p values < .05; see Table 3). Additionally, a re-administered PHQ by the research study staff also showed statistical significance (p < .05; see Table 3). However, there were no significant between-group differences on the PROMIS Global Health, PROMIS Physical Function, PROMIS Sleep Disturbance, PROMIS Satisfaction with Social Role, PROMIS Pain Interference, PROMIS Pain Intensity, CAGEAID Total, Psychiatry Consultation Index (PCI), and Charleston Comorbidity Index (CCI) as detailed in Table 3. Similar findings were evident when the PROMIS scores were converted into t-scores as depicted in Table 4. Significant between-group differences were observed for t-score mean values on the PROMIS Global Mental Health, PROMIS Anxiety, PROMIS Depression, PROMIS Fatigue, and PROMIS Cognitive Abilities measures (all p values < .05; see Table 4).

Table 3: ANOVA for Positive PHQ Screen and Negative PHQ Screen on Psychiatry Measures: Raw Mean Scores, Standard Deviations and Significance

Variable	Negative Screen		Positive Screen		*p<.05
	(n=12)		(n=16)		
	M	SD	M	SD	
PHQ Retest	14.58	7.90	19.34	5.60	*
Well-Being Index (WHO-5)	8.50	7.50	6.50	4.21	*
PROMIS GLOBAL Health	9.55	3.50	9.25	3.59	
PROMIS Global Mental Health	8.58	3.58	7.19	2.81	*
PROMIS Physical Function	10.36	6.96	10.13	5.21	
PROMIS Anxiety	14.09	4.21	15.20	3.32	*
PROMIS Depression	11.55	4.11	16.53	3.36	*
PROMIS Fatigue	14.36	5.89	17.00	2.14	*
PROMIS Sleep Disturbance	13.73	4.61	15.67	4.17	
PROMIS Satisfaction with Social	8.45	4.78	7.27	4.71	
PROMIS Pain Interference	14.73	6.17	15.53	5.28	
PROMIS Pain Intensity	6.64	3.61	6.67	2.61	
PROMIS Cognitive Abilities	10.40	5.21	8.40	4.27	*
GAGEAID Total	0.50	1.27	0.92	1.60	
Psychiatry Consultation Id	10.82	3.49	13.31	4.92	
Charleston Comorbidity Index	2.18	1.84	2.62	3.10	

Table 4: ANOVA for Positive PHQ Screen and Negative PHQ Screen on Psychiatry Measures: T-Score Mean Values, Standard Deviations and Significance

Variable	Negative Screen		Positive Screen		*p<.05
	(n=12)		(n=16)		
	M	SD	M	SD	
PROMIS GLOBAL Health	33.82	9.82	32.48	10.05	
PROMIS Global Mental Health	35.23	9.23	30.92	7.79	*
PROMIS Physical Function	34.14	12.07	37.01	12.44	
PROMIS Anxiety	66.93	8.95	70.02	7.08	*
PROMIS Depression	60.56	9.21	71.26	6.96	*
PROMIS Fatigue	62.41	14.88	67.42	5.40	*
PROMIS Sleep Disturbance	57.65	10.64	62.41	9.64	
PROMIS Satisfaction with Social	38.22	11.11	36.70	9.82	
PROMIS Pain Interference	67.26	12.58	67.55	9.05	
PROMIS Cognitive Abilities	43.34	9.40	38.83	7.46	*

There were several significant group differences on ANOVAs comparing the same clinical measures across the Depression and Non-Depression groups based on the PROMIS-Depression measure, utilizing a t-score cutoff > 65) as detailed in Table 5.

Table 5: ANOVA Comparing Depressed Versus Non-Depressed Groups Based on PROMIS Depression T-Scores > 65 for the Following Psychiatry Measures:

Variable	Non Depressed		Depressed		*p<.05
	(n=11)		(n=16)		
	M	SD	M	SD	
PHQ Retest	11.50	8.01	20.78	4.81	*
Well-Being Index (WHO-5)	12.91	6.04	04.50	3.42	*
PROMIS Global Health	12.20	3.05	7.56	2.31	*
PROMIS Global Mental Health	11.18	3.34	6.31	2.12	*
PROMIS Physical Function	12.82	6.13	8.00	4.90	*
PROMIS Anxiety	11.45	4.03	16.00	3.06	*
PROMIS Fatigue	11.73	5.68	17.81	2.04	*
PROMIS Sleep Disturbance	13.27	3.20	15.56	4.91	
PROMIS Satisfaction with Social	8.82	4.60	7.06	4.82	
PROMIS Pain Interference	12.91	5.68	16.81	5.01	
PROMIS Cognitive Abilities	11.55	6.01	8.69	4.01	
CAGEAID Total	1.09	1.58	0.57	1.28	
PCI	12.00	5.66	12.47	4.72	
CCI	3.78	3.31	1.60	1.55	*

Results demonstrated that the PHQ (both administered initially by nurses and by the research study staff), Well-Being Index, PROMIS Global Health, PROMIS Global Mental Health, PROMIS Physical Function, PROMIS Anxiety, PROMIS Fatigue, and Charleston Comorbidity Index (CCI) measures were all significantly different between groups (p < .05). Table 6 demonstrates

significant frequency differences using a Chi-square test (p < 0.05) for PHQ Depression Screen Group (Positive versus Negative) and Depressed versus Non-Depressed groups based on PROMIS Depression t-Scores > 65.

Table 6: Chi-Square for PHQ Depression Screen Group (Positive versus Negative) Versus Depressed and Non-Depressed Groups Based on PROMIS Depression T-Scores > 65

PHQ Screened by Nurse	
PROMIS Depression T-Scores > 65	
Screened Negative for Depression	
Non-Depressed (n=6)	Depressed (n=4)
Screened Positive for Depression	
Non-Depressed (n=3)	Depressed (n=12)

*Pearson Chi-Square p<.05

Lastly, a correlational matrix between psychiatry clinical measures indicated that the PHQ-Nurse indicated significant correlations with the PROMIS Physical Function, PROMIS Depression, PROMIS Sleep Disturbance, and PROMIS Pain Interference measures. PHQ retest data indicated correlations with the WHO-5, PROMIS Global Health, PROMIS Mental Health, PROMIS Anxiety, PROMIS Depression, and PROMIS Pain Intensity measures. Detailed correlation results are provided in Table 7.

Table 7: Correlation Matrix for PHQ Nurse Screen and PHQ Retest with other Psychiatry Measures

	PHQ Nurse	PHQ Retest
	(n=16)	(n=27)
PHQ Nurse	N/A	.249
PHQ Retest	.249	N/A
Well-Being Index (WHO-5)	-.481	-.390*
PROMIS GLOBAL Health	-.318	-.374*
PROMIS Global Mental Health	-.353	-.439*
PROMIS Physical Function	-.557*	-.003
PROMIS Anxiety	.149	.453*
PROMIS Depression	.636*	.600*
PROMIS Fatigue	.291	.333
PROMIS Sleep Disturbance	.559*	.187
PROMIS Satisfaction with Social	-.129	.069
PROMIS Pain Interference	.554*	.177
PROMIS Pain Intensity	.555*	.414*
PROMIS Cognitive Abilities	.268	-.326
GAGEAID Total	-.302	-.062
Psychiatry Consultation (PCI)	-.425	-.301
Charleston Comorbidity Index	-.359	-.352

*p<.05

4. Discussion

As comorbid medical and psychiatric illness continue to go unrecognized in academic medical centers, the importance of screening for psychiatric illness, particularly depression, during the admission process for inpatient care becomes increasingly important. While the need for such screening services has valid and empirical support (as evidenced by our findings), there continues to be few medical centers with existing infrastructure that screen for depression. We were particularly concerned with depression, given the chronic and persistent course of the illness, and relatively high prevalence rates (approximately 10-30% of patients), all of which frequently result in disabling illness that can ultimately lead to increased mortality and morbidity. As a consequence, there are greater costs in healthcare systems that result from failure to identify underlying psychiatric illness, which ultimately leads to worse health outcomes. The Cedars-Sinai Medical Center collaborative, interdisciplinary, and comprehensive psychiatric screening service involves identifying patients who screen positive for depression

for the purpose of informing their treating physician, obtaining social work consult, and if needed, referring them to the existing consultation-liaison team housed in the Department of Psychiatry and Behavioral Neurosciences. The specific goals of developing this program were to systematically and proactively identify medical-surgical inpatients with substantial underlying depression. In doing so, we hope to increasingly become able to examine the relationships between psychiatric symptoms/disorders, treatments on medical conditions, patient satisfaction, QOL, functioning, readmission rates, length of stay, and overall health in a large academic medical center. Furthermore, data collected for the study will help inform our understanding of the most sensitive, efficient, and appropriate measures to assess these various constructs. This will ultimately increase time and effectiveness in resource limited settings. The novelty of our research design is continually allowing us to gather "normative data" for inpatient populations across a wide variety of patient self-reported measures of QOL, functioning, overall well-being, health, and psychiatric symptom severity.

Prior research has shown the validity and utility of screening measures for depression in various settings. A meta-analysis including 17 validation studies with a total population of 5,000 showed that for MDD not only did the PHQ-9 have good diagnostic properties (92% sensitivity and 80% specificity), but that the diagnostic properties were relatively consistent among a range of settings and specialties including community, primary care, and various hospital specialties. In addition, research has shown the PHQ-9 to work well in a variety of cultures and with a variety of translations which is comparable to or better than clinician administered instruments for MDD screening (Gilbody et al. 2006). Furthermore, in a study of around 200 patients admitted over a 6 month period to internal medicine departments of hospitals who filled out a PHQ-9 and had a psychiatric evaluation, found that although the PHQ-9 was 5.7% more likely to overestimate minor depression versus psychiatric evaluation and 3.8% more likely to associate MDD with a somatic disorder, they are needed as only 19% of moderately severe depressions and 18.7% of severe depressions received a diagnosis of depression from the internists and even fewer (9.5% and 12.5%, respectively) had been treated by internists with antidepressants (Rentsch et al. 2007).

Research has identified the importance of interdisciplinary and multidisciplinary teams, as being useful in providing comprehensive and collaborative care in medical settings, given the increased benefit of the combined efforts of physicians, psychologist, social workers, and other specialist; all of which can have a significantly positive effect on the delivery of care, with the ultimate goal of increasing long-term health outcomes for our patients (Winters & Metz 2009).

One of the most significant contributions in the current study was the inclusion of scores for patients in an inpatient medical setting across a wide range of frequently used measures and the correlation matrix that provided test-retest reliability for the selected measures. The implications of these findings warrant continued investigation of these relationships utilizing a larger sample size. However, we are still in the initial stages of establishing the data-registry that we hope will eventually result in a large and nationally representative sample of participants' scores while receiving inpatient care in a large medical center. The ultimate goal is to continue to identify the risk factors associated with poor health outcomes and compromised well-being, including but not limited to failure to achieve remission and reentry into the hospital setting. In addition, continued data collection will ultimately help us identify the best measures or instruments to administer, resulting in greater sensitivity and specificity as well as greater positive and negative predictive values and accuracy to identify at risk groups.

Study Limitations and Strengths

In establishing a depression screening service like the one proposed in this study, there are costs associated with such a program existing including the burden of time and effort for staff and patients. The risk of false positives and false negatives can also occur. However, the benefits outweigh the costs such that if a patient is screened for depression, it is better than not receiving such a

screening. In the cases of a false positive, a patient is then referred to the consultation-liaison psychiatry team where they can be further assessed, and if it is deemed that the initial screening was in fact a false positive, the worst consequence is some expended time and effort. The alternative however, is far worse and could result in failure to identify, diagnose, and prevent further suffering. Another limitation in the current study includes the relatively small sample size. While this may be a concern, this type of pilot information is valuable and there is a plan to continue gathering data and increase the sample size to establish medically ill inpatient normative scores for all the patient-reported outcome measures included in our study. Further limitations of this study include the lack of data on reading levels, translated versions, and the efficacy of the instruments in different cultures especially given the non-English speaking population of Los Angeles and the US. Lastly, the study could have been improved by implementing a double or single blind design.

Strengths of the current study include the wide age range of participants as well as a balanced ratio of male and female participants, in addition to a relatively diverse ethnic sample that was representative of the greater Los Angeles area. However, there are plans to include wider ranges of ethnic, racial, and underrepresented groups through the continued duration of the study, as doing so will strengthen the generalizability of our findings. Some practicing clinicians as well as researchers have critiqued the use or reliance upon patient-reported measures as valid indices of objective psychological and emotional functioning. However, every measure included in the current study has a significant amount of support for strong psychometric properties, including robust validity, and reliability for the constructs being assessed. Furthermore, there has been a shift in which the United States Food and Drug Administration, World Health Organization, and National Institutes of Health Patient Reported Outcomes Measurement Information System (PROMIS) initiative all support the use of patient-reported measures in psychiatric, medical, and clinical work as well as research pursuits.

5. Conclusion

Findings from the current study are promising and provide support for the clinical utility of establishing a psychiatric screening service for depression during inpatient hospital admissions (Ebell 2008). As other academic medical centers and teaching hospitals adopt a model similar to the one in the current study, the results will prove to be cost-effective. Furthermore, the importance of screening for depression is quite evident, given the fatal result of completed suicide and increased risk for suicide attempts (Angst et al. 1999, Osvath et al. 2004). In addition, our findings demonstrated significantly worse QOL, health, functioning, and psychiatric symptom severity for depressed patients compared to non-depressed patients. Not surprisingly, our suggestions are supported by the National Patient Safety Goal 15.01.01 that was established by The Joint Commission, which requires behavioral health care organizations, and psychiatric and general hospitals that treat individuals for emotional and/or behavioral disorders to identify individuals at risk for suicide (The Joint Commission 2016). Hospitals that establish a system similar to ours, act as gatekeepers to an individual's possibility and potential for recovery, where failure to do so can have devastating consequences.

References

[1] Aertgeerts B, Buntinx F, Kester A (2004) the value of the CAGE in screening for alcohol abuse and alcohol dependence in general clinical populations: a diagnostic meta-analysis. Journal of Clinical Epidemiology 57, 30-9. http://dx.doi.org/10.1016/S0895-4356(03)00254-3.

[2] American Psychiatric Association (APA) (2000). Diagnostic and Statistical Manual of Mental Disorders 4th edition text revision. Washington, DC: American Psychiatric Association. http://dx.doi.org/10.1176/appi.books.9780890423349.

[3] Angst J, Angst F, Stassen HH (1999) Suicide risk in patients with major depressive disorder. The Journal of Clinical Psychiatry 60, 57-62; discussion 75-6, 113-6.

[4] Armitage GD, Suter E, Oelke ND, Adair CE (2009) Health systems integration: state of the evidence. International Journal of Integrated Care 9, e82. http://dx.doi.org/10.5334/ijic.316.

[5] Bech P (2004). Measuring the dimension of psychological general well-being by the WHO-5. Quality of Life Newsletter 20, 15-6.

[6] Bech P (2012) Clinical psychometrics. John Wiley & Sons. http://dx.doi.org/10.1002/9781118511800.

[7] Bech P, Olsen LR, Kjoller M, Rasmussen NK (2003) Measuring well-being rather than the absence of distress symptoms: a comparison of the SF-36 Mental Health subscale and the WHO-Five Well-Being Scale. International Journal of Methods in Psychiatric Research 12, 85-91. http://dx.doi.org/10.1002/mpr.145.

[8] Bonicatto SC, Dew MA, Zaratiegui R, Lorenzo L, Pecina P (2001) Adult outpatients with depression: worse quality of life than in other chronic medical diseases in Argentina. Social Science and Medicine 52, 911-9. http://dx.doi.org/10.1016/S0277-9536(00)00192-1.

[9] Cella D, Riley W, Stone a, Rothrock N, Reeve B, Yount S (2010) The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. Journal of Clinical Epidemiology 63, 1179-94. http://dx.doi.org/10.1016/j.jclinepi.2010.04.011.

[10] Centers for Disease Control (CDC) Chronic disease and mental health. Available at: http://www.cdc.gov/mentalhealth/information/chronic-disease.htm. Accessed June 12 2016.

[11] Chapman DP, Perry GS, Strine TW (2005) the Vital Link between Chronic Disease and Depressive Disorders. Preventing Chronic Disease 2, A14.

[12] Charlson ME, Pompei P, Ales KL, MacKenzie CR (1987) a new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of Chronic Diseases, 40: 373-83. http://dx.doi.org/10.1016/0021-9681(87)90171-8.

[13] Dhalla S, Kopec JA (2007) The CAGE questionnaire for alcohol misuse: a review of reliability and validity studies. Clinical and Investigative medicine 30, 33-41.

[14] Ebell MH (2008) Screening instruments for depression. American Family Physician 78,244-6.

[15] El-Gabalawy R, Katz LY, Sareen J (2010) Comorbidity and associated severity of borderline personality disorder and physical health conditions in a nationally representative sample. Psychosomatic Medicine 72, 641-7. http://dx.doi.org/10.1097/PSY.0b013e3181e10c7b.

[16] Enthoven AC (2009) integrated delivery systems: the cure for fragmentation. The American Journal of Managed Care 15, S284-90.

[17] Evans DL, Charney DS, Lewis L, Golden RN, Gorman JM, Krishnan KR (2005) Mood disorders in the medically ill: scientific review and recommendations. Biological Psychiatry 58, 175-89. http://dx.doi.org/10.1016/j.biopsych.2005.05.001.

[18] Ewing JA. Detecting alcoholism (1984) The CAGE questionnaire. JAMA 252, 1905-7. http://dx.doi.org/10.1001/jama.1984.03350140051025.

[19] Gilboy S, Bower P, Fletcher J, Richards D, Sutton AJ (2006) Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. Archives of Internal Medicine 166, 2314-21. http://dx.doi.org/10.1001/archinte.166.21.2314.

[20] Hays R, Bjorner J, Revicki D, Spritzer K, Cella D (2009) Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Quality of Life Research 18,873-80. http://dx.doi.org/10.1007/s11136-009-9496-9.

[21] Henkel V, Mergl R, Kohnen R, Maier W, Möller H-J, Hegerl U (2003) Identifying depression in primary care: a comparison of different methods in a prospective cohort study. British Medical Journal 326,200-1. http://dx.doi.org/10.1136/bmj.326.7382.200.

[22] IsHak WW, Becker B, Hedrick R, Solt A, Lopez E, Peterson M, Danovitch I (2014) Proactive Consultations in a General Hospital using the Cedars-Sinai Psychiatric Consultation Identification Score (CS-PCI). Oral presentation presented at the 65th Annual Meeting of The Academic of Psychosomatic Medicine, Fort Lauderdale, FL.

[23] Katon W (1996) the impact of major depression on chronic medical illness. General Hospital Psychiatry 18, 215-9. http://dx.doi.org/10.1016/0163-8343(96)00065-5.

[24] Kessler RC, Heeringa S, Lakoma MD, Petukhova M, Rupp AE, Schoenbaum M (2008) Individual and societal effects of mental disorders on earnings in the United States: results from the national

comorbidity survey replication. American Journal of Psychiatry 165, 703-11. http://dx.doi.org/10.1176/appi.ajp.2008.08010126.

[25] Kroenke K, Spitzer RL, Williams JB (2001) the PHQ-9: Validity of a Brief Depression Severity Measure. Journal of General Internal Medicine 16,606-13. http://dx.doi.org/10.1046/j.1525-1497.2001.016009606.x.

[26] Kroenke K, Spitzer RL, Williams JB (2003) The Patient Health Questionnaire-2: validity of a two-item depression screener. Medical Care 41, 1284-92. http://dx.doi.org/10.1097/01.MLR.0000093487.78664.3C.

[27] National Institute of Mental Health (NIMH) Any Mental Illness (AMI) Among Adults. Available at: http://www.nimh.nih.gov/health/statistics/prevalence/any-mental-illness-ami-among-us-adults.shtml. Accessed June 12 2016.

[28] Osvath P, Voros V, Fekete S (2004) Life events and psychopathology in a group of suicide attempters. Psychopathology 37, 36-40. http://dx.doi.org/10.1159/000077018.

[29] Rentsch D, Dumont P, Borgacci S, Carballeira Y, deTonnac N, Archinard M, Andreoli A (2007) Prevalence and treatment of depression in a hospital department of internal medicine. General Hospital Psychiatry 29, 25-31. http://dx.doi.org/10.1016/j.genhosppsych.2006.08.008.

[30] Saffer BY, Lanting SC, Koehle MS, Klonsky ED, Iverson GL (2015). Assessing cognitive impairment using PROMIS ((R)) applied cognition-abilities scales in a medical outpatient sample. Psychiatry Research 226,169-72. http://dx.doi.org/10.1016/j.psychres.2014.12.043.

[31] The Joint Commission (2016) National Patient Safety Goals. Available at http://www.jointcommission.org/standards_information/npsgs.aspx. Accessed June 12, 2016.

[32] Üstün TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J (2010) Developing the World Health Organization Disability Assessment Schedule 2.0. Bulletin of the World Health Organization 88, 815-23. http://dx.doi.org/10.2471/BLT.09.067231.

[33] Verbosky LA, Franco KN, Zrull JP (1993) the relationship between depression and length of stay in the general hospital patient. The Journal of Clinical Psychiatry 54,177-81.

[34] Winters NC, Metz WP (2009) the wraparound approach in systems of care. The Psychiatric Clinics of North America 32,135-51. http://dx.doi.org/10.1016/j.psc.2008.11.007.

[35] World Health Organization (WHO) (1998). Wellbeing measures in primary health care: the DepCare project: report on a WHO meeting Stockholm, Sweden 12-13 February 1998. WHO Regional Office for Europe.