



“Thandi should feel embarrassed”: describing the validity and reliability of a tool to measure depression-related stigma among patients with depressive symptoms in Malawi

Josée M. Dussault¹ · Chifundo Zimba² · Jullita Malava³ · Harriet Akello² · Melissa A. Stockton^{4,5} · Michael Udedi⁶ · Bradley N. Gaynes^{1,7} · Mina C. Hosseinipour^{2,8} · Brian W. Pence¹ · Jones Masiye⁶

Received: 7 April 2021 / Accepted: 10 November 2021 / Published online: 20 November 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany 2021

Abstract

Purpose There are no validated tools in Malawi to measure mental health stigma. Accordingly, this study evaluates the validity and reliability of a short quantitative instrument to measure depression-related stigma in patients exhibiting depressive symptoms in Malawi.

Methods The SHARP study began depression screening in 10 NCD clinics across Malawi in April 2019; recruitment is ongoing. Eligible participants were 18–65 years, had a patient health questionnaire (PHQ-9) score ≥ 5 , and were new or current diabetes or hypertension patients. Participants completed a baseline questionnaire that measured depression-related stigma, depressive symptoms, and sociodemographic information. The stigma instrument included a vignette of a depressed woman named Thandi, and participants rated their level of agreement with statements about Thandi’s situation in nine prompts on a 5-point Likert scale. Inter-item reliability was assessed with Cronbach’s alpha. Exploratory factor analysis (EFA) was used to assess structural validity, and OLS regression models were used to assess convergent and divergent validity between measured levels of depression-related stigma and covariates.

Results The analysis of patient responses ($n = 688$) to the stigma tool demonstrated acceptable inter-item reliability across all scales and subsequent subscales of the instrument, with alpha values ranging from 0.70 to 0.87. The EFA demonstrated clustering around three domains: negative affect, treatment carryover, and disclosure carryover. Regression models demonstrated convergence with several covariates and demonstrated divergence as expected.

Conclusion This study supports the reliability and validity of a short stigma questionnaire in this population. Future studies should continue to assess the validity of this stigma instrument in this population.

Keywords Mental health stigma · Depression stigma · Malawi

✉ Josée M. Dussault
josee.dussault@unc.edu

¹ Department of Epidemiology, UNC Gillings School of Global Public Health, Chapel Hill, USA

² UNC Project, Malawi, Lilongwe, Malawi

³ Malawi Epidemiology and Intervention Research Unit, Karonga, Malawi

⁴ Department of Psychiatry, Columbia University Vagelos College of Physicians and Surgeons, New York, USA

⁵ New York State Psychiatric Institute, New York, USA

⁶ NCDs and Mental Health Unit, Ministry of Health Malawi, Lilongwe, Malawi

⁷ Department of Psychiatry, UNC School of Medicine, Chapel Hill, USA

⁸ Division of Infectious Disease, UNC School of Medicine, Chapel Hill, USA

Background

Globally, mental illness is prevalent. The 2017 Global Burden of Disease (GBD) study estimated that 970 million people—approximately 13% of the global population—were living with a mental disorder [1, 2]. Approximately 264 million people, or 4% of the global population, were estimated to be living with a depressive disorder [2]. Although there are cost-efficient and effective treatment options to address depression [3, 4], large gaps exist in care-seeking and receipt of effective treatment. One major barrier affecting care-seeking among patients exhibiting depressive symptoms is stigma, the deeply entrenched sociocultural norms that comprise attitudes which reduce people to being tainted or otherwise discounted based on some attribute [5, 6]. Since what is considered to be a discreditable attribute—and the degree to which it discredits—may vary by cultural context, tools measuring stigma may need to be tailored to and validated in specific populations.

In European and North American countries, where the bulk of mental health stigma research has been based, those who report greater mental health stigmatization are less likely to seek care and are more likely to experience discrimination and demoralization [7–10]. Beyond seeking treatment, stigma also negatively affects adherence and response to treatment [11, 12]. Previous studies have also demonstrated that the experiences and consequences of stigma are modified by other social attributes of patients, including age, gender, education, and socioeconomic status, demonstrating that the social context of mental health stigma is very important [13–16].

In many African countries, there has yet to be extensive research on the role of mental health stigma in mental health treatment. One study conducted in South Africa found that stigma and misinformation regarding mental illness are prevalent and negatively associated with help-seeking behavior [17]. A cross-sectional study of attitudes around mental illness in Blantyre, Malawi observed that the degree and type of stigmatizing beliefs that participants endorsed varied by age but not by education, socioeconomic status, or gender [18]. While that study produced valuable information, it did not use an instrument validated in the study population, as there have yet to be such tools validated in Malawi to measure mental health stigma.

Accordingly, this study aims to evaluate the structural validity and reliability of a short quantitative instrument to measure depression-related stigma in patients engaged in noncommunicable disease (NCD) care and exhibiting depressive symptoms in Malawi. We further aim to characterize the prevalence of depression-related stigma among

this cadre of patients and assess construct validity, or the degree to which this measure of stigma is associated with other measures of other constructs as we would expect [19].

Methods

Study design

The Sub-Saharan Africa Regional Partnership for Mental Health Capacity Building (SHARP) began depression screening in 10 NCD clinics in Malawi in May 2019, and recruitment to the cohort is ongoing. The overall objective of the SHARP scale-up study is to integrate depression treatment with diabetes and hypertension treatment at NCD clinics and compare outcomes between basic and enhanced clinic-level implementation strategies. Patients were recruited into the SHARP cohort by consecutive depression screening as they presented to the NCD clinics for their standard care. Eligible participants were 18–65 years of age, had elevated depressive symptoms denoted by a score ≥ 5 on the Patient Health Questionnaire (PHQ-9) [20–23], and had a new or current diagnosis of diabetes or hypertension. SHARP research assistants asked participants a series of questions, in Chichewa or Chitumbuka, to complete the baseline questionnaire. Patients included in this cross-sectional analysis completed their baseline interviews between May 2019 and March 2021.

Outcome assessment: depression-related stigma

The outcome of interest, patients' levels of depression-related stigma, was measured using a brief 10-item instrument that was adapted from the Stigma in Global Context—Mental Health Study (SGC-MHS) [24, 25]. The SHARP study team used an iterative approach to tailor the instrument to the social and linguistic context of the study: team members involved in the study's protocol development worked together to identify key prompts from the original SGC-MHS instrument to adopt and translate to Chichewa and Chitumbuka in a way that maintained the meaning of the prompt in English while being understandable in the target languages and cultures. Although the original SGC-MHS stigma instrument was 75 items, the SHARP team generated a modified 10-item instrument to meet the SHARP study's data collection needs while focusing on facets of stigma that the team anticipated would be salient in the Malawian context based on prior research experience around depression in Malawi [26–28]. Ultimately, questions in the final instrument were related to negative feelings and attitudes toward individuals with depression (negative affect), the role of disclosure particularly on the family (disclosure carryover), and

social isolation as a result of engaging in treatment (treatment carryover) [29, 30]. Disclosure carryover prompts were centered around the family due to the importance of family in this population and the indication from previous research that stigma spills over onto the family [18, 31]. Study team members who translated the stigma instrument were fluent in the respective target language.

The stigma instrument first introduced a vignette of a woman named Thandi who was experiencing depressive symptoms, without naming it as depression. Participants were asked to rate their level of agreement with statements about whether Thandi's situation was shameful or embarrassing, e.g., in four prompts on a 5-point Likert scale from strongly disagree to strongly agree. A second vignette then explained that Thandi had been diagnosed with depression and presented five more prompts, almost parallel to the first four prompts, on the same 5-point Likert scale. Each prompt was written such that agreement endorsed depression-related stigma. Strong agreement with any given prompt was equivalent to 4 points, while strong disagreement was 0 points.

Along with the nine Likert scale items in the stigma tool, a tenth item asked participants to identify, from a list of options, what they believed to be the most likely cause of Thandi's behavior (prior to the second vignette). Participants had seven prespecified responses to choose from, but they were also able to provide another response not listed or respond that they did not know. Responses to this item were not quantified but were instead examined graphically to better understand the way that participants interpreted Thandi's depressive symptoms. If not belonging to the seven prespecified etiologies, patient responses were included in the "other" category and not further described.

Based on the exploratory factor analysis (EFA) results, stigma subscales were produced as the average of the items belonging to each factor identified. The overall stigma score was then equivalent to a weighted average of the subscale measures, based on the number of questions each subscale provided to the overall scale. To describe study sample characteristics succinctly, participants with an overall stigma score greater than 2 were identified as having high stigma. Using the Likert scale described previously (ranging 0–4, with 2 corresponding to "Neutral"), participants with an overall stigma score greater than 2 on average expressed agreement with the statements articulating stigmatizing beliefs, whereas those with a score of 2 or less were on average neutral or expressed disagreement with the statements.

Covariate assessment

Patient baseline interviews included instruments that measured depression (PHQ-9) [20], stressful life events (Life Events Scale) [32, 33], social support (Multidimensional Scale of Perceived Social Support) [34, 35], and adaptive

coping behaviors (Brief COPE) [36]. Adaptive coping behaviors included positive reframing, acceptance, humor, religion, and using emotional support [36]. Baseline demographic information was also included, such as sex (as a proxy for gender), age, education level, and indicators of socioeconomic status. A weighted factor score was used as an indicator of socioeconomic status, generated from responses to radio, refrigerator, television, mobile phone, and car ownership, along with responses to whether the household had electricity, how often they worried about money, and how often they went to bed hungry. All scales were coded such that a larger number equated to a greater level or amount of that construct (e.g., greater wealth than the average participant or greater social support). Because the wealth score was generated as a weighted factor score, the mean will be expected to be equal to 0 and the standard deviation will be equal to 1 in this sample.

Statistical methods

To begin understanding the reliability and validity of the stigma measure, the distribution of characteristics of the SHARP baseline cohort was described in a table stratified by high versus low stigma. A Chi-square test was used to compare participants' ability to correctly identify Thandi's depressive symptoms across the high versus low stigma categories. Inter-item reliability was evaluated using Cronbach's alpha. We then assessed three types of validity: structural, convergent, and divergent. To assess structural validity, an exploratory factor analysis (EFA) was conducted using the principal-factors method with a Promax rotation. Scree plots were used to identify the proper number of factors to retain in the analysis prior to rotation [37]. Using an EFA allowed us to investigate the dimensionality of the stigma tool and identify problematic items [38]. To assess convergent validity in this patient population, OLS regression models were run that included expected correlates of depression-related stigma (age, sex, education, depressive symptoms, social support, stressful life events, coping behaviors, and wealth). To identify divergent validity, a separate OLS regression model was run, regressing stigma scores onto height (measured in centimeters), while adjusting for sex. The hypothesis was that height and stigma would be unassociated, while stigma levels may be related to patients' age, depression severity, and other indicators of social status. Stigma scores, depressive symptoms, age, social support, stressful life events, coping behaviors, and wealth scores were all included in models as continuous variables, whereas sex was coded as a binary categorical variable with female as the reference level, and education level was a 5-level ordinal categorical variable with no formal schooling as the reference level. Given that few datapoints were missing (3% in convergent validity analysis; 9% in divergent validity analysis),

complete case analyses were conducted. All continuous variables were assessed for best functional form using marginal graphical analyses and likelihood ratio tests, which demonstrated that all continuous variables in these regression models fit best as simple linear terms. The data were analyzed using Stata [39].

Results

Of the 1586 patients approached by research staff for eligibility assessment at the time of data extraction for this analysis, 1105 were eligible and 695 consented. Ultimately, 689 participants completed their baseline interviews, and 688 participants completed the stigma questionnaire during their baseline interviews. Among these participants, 338 (49%) were identified as having high depression-related stigma (Table 1). The sample included 556 (81%) female patients

and 132 (19%) male patients, and 380 (55%) of the patients admitted to often worrying about not having enough money for basic necessities. The total sample had overall stigma scores that were normally distributed (mean: 2.1; SD: 0.8).

The stigma instrument demonstrated acceptable inter-item reliability with a Cronbach's alpha of 0.77 (Table 2). Scree plots produced from the EFA suggested that there were three factors within the stigma instrument, and the subsequent rotated factor loadings clustered around three domains of stigma (negative affect, disclosure carryover, and treatment carryover) [29]. The EFA results demonstrated high levels of agreement among the questions that belong to each domain, with the exception of item #4 (Table 2). Thus, item #4 was removed in a second factor analysis and was not included in the scoring of stigma scales and subscales in this analysis. The removal of item #4 did not dramatically change subscale alpha values or the factor loadings of other items in the stigma instrument (see Online Resource 1). Within

Table 1 Characteristics of SHARP patients by baseline depression stigma ($N=688$)

	High stigma score (> 2; $n=338$)		Low stigma score (≤ 2 ; $n=350$)		Total	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Education level						
No formal schooling	66	20%	55	16%	121	18%
Standard 1–5	122	36%	115	33%	237	34%
Standard 6–8	89	26%	109	31%	198	29%
Secondary school	44	13%	58	17%	102	15%
Post-secondary school	17	5%	13	4%	30	4%
Sex						
Female	276	82%	280	80%	556	81%
Male	62	18%	70	20%	132	19%
Worried about money						
Often	194	57%	186	53%	380	55%
Sometimes	98	29%	119	34%	217	32%
Never	46	14%	45	13%	91	13%
	Mean	SD	Mean	SD	Mean	SD
Age	51.0	9.5	50.5	9.7	50.7	9.6
Number of people in household	3.3	1.7	3.1	1.6	3.2	1.6
PHQ-9 score (0–27)	8.6	4.9	7.7	4.0	8.2	4.5
GAD-7 score (0–27)	6.7	4.6	6.1	4.0	6.4	4.3
Social support score (MSPSS; 0–24)	17.2	4.0	17.2	4.1	17.2	4.0
Stressful life events (LES; 0–29)	2.8	2.4	2.8	2.4	2.8	2.4
Adaptive behaviors (0–15)	6.7	3.0	7.7	2.8	7.2	3.0
Wealth score (– 3–3) ^a	0.0	1.0	0.0	1.0	0.0	1.0
Overall stigma scale	2.7	0.4	1.5	0.5	2.1	0.8
Negative affect subscale	2.8	0.6	1.6	0.8	2.2	0.9
Disclosure carryover subscale	3.0	0.7	1.9	1.2	2.4	1.1
Treatment carryover subscale	2.1	1.2	0.8	0.6	1.5	1.1

Wealth score was generated as a weighted factor score; therefore, the mean is equal to 0 and the standard deviation is equal to 1 in this sample. The range of possible values is then – 3 to +3 as a feature of the factor score, too.

Table 2 Reliability and structural validity of the overall stigma scale and subscales

Questionnaire items	Factor 1: negative affect	Factor 2: treatment car- ryover	Factor 3: disclosure car- ryover	Alpha Value
Overall stigma scale				0.77
Negative affect subscale				
1. Being around Thandi would make me feel uncomfortable	0.61	– 0.03	– 0.09	0.70
2. Thandi should feel embarrassed about her situation	0.51	– 0.18	0.22	
4. If people found out about Thandi’s situation, Thandi would lose some of her friends	0.29	0.25	0.03	
6. Being around someone diagnosed with depression like Thandi would make me feel uncomfortable	0.66	0.10	– 0.10	
7. Thandi should feel embarrassed about her diagnosis of depression	0.52	0.07	0.14	
Disclosure carryover subscale				
3. Members of Thandi’s family would be better off if Thandi’s situation was kept secret	– 0.08	0.02	0.69	0.74
8. Members of Thandi’s family would be better off if her diagnosis of depression was kept secret	0.04	0.05	0.68	
Treatment carryover subscale				
9. If Thandi were going to a clinic regularly to speak with a trained therapist about her problems to help treat her depression, and people found out she was going, she would lose some of her friends	0.01	0.82	0.03	0.87
10. If Thandi were taking medication to help treat her depression, and people found out she was taking medication, she would lose some of her friends	– 0.01	0.82	0.00	

each factor, the Cronbach’s alpha remained acceptable, demonstrating sufficient internal consistency within subscales of the stigma instrument (Table 2). Lastly, the correlation matrix in Table 3 demonstrates the distinct nature of each domain, with low correlation coefficients between subscales.

In identifying Thandi’s condition in the initial vignette, prior to being told that Thandi had been diagnosed with depression, 36% of participants said she had depression, while 34% thought it was stress, and 7% believed it was the normal ups and downs of life. Other responses were poverty (6%), HIV (5%), asthma (1%), and schizophrenia (2%). Another 6% of participants said they did not know (see Online Resource 2). Among participants with high stigma (overall score > 2), 114 (34%) identified Thandi’s condition as depression as compared to 137 (39%) patients in the low (overall score ≤ 2) stigma group. In contrast, 128 (38%) patients in the high stigma group identified her condition as stress, as compared to 103 (29%) patients in the low stigma group ($\chi^2 = 12.7, df = 7, p = 0.08$).

Generally, patients in the high stigma and low stigma groups had similar characteristics (Table 1). Patients with high and low stigma scores shared comparable distributions of education levels. Both groups also had similar distributions across age, sex, and socioeconomic status. On average, the high stigma group had greater depressive and anxiety symptoms, while the low stigma group had greater social support and more adaptive behaviors (Table 1).

The results of the convergent validity analysis demonstrated varying patterns of association between the four stigma scales—the overall scale and the three subscales—and the set of covariates. For example, regression coefficients from Model 1—with the overall stigma score as the dependent variable—indicated that greater depressive symptoms were positively associated with the overall stigma score, while adaptive coping behaviors were negatively associated with the overall stigma score (Table 4). Model 2—describing disclosure carryover—found that depressive symptoms and social support were positively associated with

Table 3 Correlations between stigma subscales

	Negative affect subscale	Disclosure carryover subscale	Treatment carryover subscale
Negative affect subscale	1.00		
Disclosure carryover subscale	0.32	1.00	
Treatment carryover subscale	0.35	0.28	1.00

Table 4 Convergent validity: coefficients and 95% CIs from ordinary least squares regression models relating the overall stigma scale and each subscale to demographic and psychosocial factors ($n=671$)^a

Independent Variables	Dependent variable			
	Model 1: overall stigma scale	Model 2: disclosure carryover subscale	Model 3: negative affect subscale	Model 4: treatment carryover subscale
	Coefficient (95% CI)			
Sex (Ref: Female)	- 0.04 (- 0.2, 0.1)	- 0.08 (- 0.3, 0.2)	- 0.10 (- 0.3, 0.1)	0.11 (- 0.1, 0.3)
Age, per 10 years	- 0.03 (- 0.1, 0.0)	- 0.09 (- 0.2, 0.0)	- 0.05 (- 0.1, 0.0)	0.07 (0.0, 0.2)
Education level (Ref: No formal school)	- 0.06 (- 0.1, 0.0)	- 0.07 (- 0.2, 0.0)	- 0.07 (- 0.2, 0.0)	- 0.01 (- 0.1, 0.1)
PHQ-9 score, per 1 unit	0.02 (0.0, 0.0)	0.03 (0.0, 0.1)	0.02 (0.0, 0.0)	0.02 (0.0, 0.0)
Wealth score, per 1 standard deviation	- 0.02 (- 0.1, 0.0)	- 0.09 (- 0.2, 0.0)	0.03 (0.0, 0.1)	- 0.05 (- 0.1, 0.0)
Stressful life events, per event	- 0.01 (0.0, 0.0)	- 0.06 (- 0.1, -0.0)	0.00 (0.0, 0.0)	0.04 (0.0, 0.1)
Social support score, per 1 unit	0.01 (0.0, 0.0)	0.03 (0.0, 0.0)	0.01 (0.0, 0.0)	- 0.01 (0.0, 0.0)
Adaptive coping score, per 1 unit	- 0.06 (- 0.1, 0.0)	- 0.04 (- 0.1, 0.0)	- 0.04 (- 0.1, 0.0)	- 0.11 (- 0.1, - 0.1)

^aBolded estimates have 95% confidence intervals that exclude the null value of zero. This may not always be clear due to rounding of small values. Analysis sample was $n=671$ due to missing data for $n=17$ participants

an increased belief that Thandi's family would be better off if her situation were kept secret. Model 2 also found that age, stressful life events, and adaptive coping behaviors were all negatively associated with the disclosure subscale (Table 4). Model 3 found that depressive symptoms were positively associated with negative affect toward Thandi; adaptive coping behaviors were negatively associated with negative affect (Table 4). Model 4 found that stressful life events were positively associated with concerns that Thandi would lose friends if she sought treatment and others found out (treatment carryover). Model 4 also found a negative association between adaptive coping behaviors and treatment carryover concerns (Table 4). In Models 5–8, which assessed divergent validity by regressing each scale onto height, adjusting for sex, the mean estimates demonstrated that a 10-cm change in height was not associated with any significant change in stigma scores, as mean estimates were centered around zero, with confidence intervals on either side of the null value of zero (Table 5; Models 5–8).

Discussion

This study describes depression-related stigma among patients with diabetes or hypertension and depressive symptoms in Malawi. The results of this study demonstrate that a large proportion of these patients (49%), on average, agreed with stigmatizing statements about individuals with depression and their anticipated experiences. Results from this study also support the reliability and validity of a short stigma questionnaire in this population. When this questionnaire was drafted, the study team used an iterative process to identify key prompts to include in the abbreviated tool. These prompts were grouped around three themes: negative feelings and attitudes toward individuals with depression (negative affect), the role of disclosure particularly on the family (disclosure carryover), and social isolation as a result of engaging in treatment (treatment carryover). The stigma tool was also designed to have parallel questions related to these themes before and after the second vignette revealed

Table 5 Divergent validity: coefficients and 95% CIs from ordinary least squares regression models relating the overall stigma scale and each subscale to height, adjusting for sex ($n=624$)^a

Independent variables	Dependent variable			
	Model 5: overall stigma scale	Model 6: disclosure carryover subscale	Model 7: negative affect subscale	Model 8: treatment carryover subscale
	Coefficient (95% CI)			
Sex (Ref: Female)	- 0.09 (- 0.3, 0.1)	- 0.10 (- 0.4, 0.2)	- 0.16 (- 0.4, 0.1)	0.08 (- 0.2, 0.3)
Height, per 10 cm	- 0.01 (- 0.1, 0.1)	0.05 (- 0.1, 0.2)	- 0.02 (- 0.1, 0.1)	- 0.05 (- 0.2, 0.1)

^aMore patients are missing data for height ($n=54$), as compared to other measured variables, resulting in a smaller sample size in this divergent validity analysis as compared to the convergent validity analysis. In total, 9% of the study sample had missing data, and complete case analysis was therefore still appropriate

Thandi's depression diagnosis with the goal of understanding the role of an official diagnosis on stigmatizing beliefs. Exploratory factor analysis demonstrated that the items grouped around the three themes without regard to whether the prompt was delivered before or after the second vignette, meaning that participants generally responded with similar levels of stigma in these three domains, regardless of whether they knew that Thandi was diagnosed with depression.

The exploratory factor analysis results also demonstrated that the fourth item of the instrument was weakly associated with two domains, likely due to poor wording of the prompt. The fourth item was thus excluded from the stigma scale and subscale scoring. The exploratory factor analysis also found that each item of the newly 8-item quantitative instrument mapped onto only one of the three domains, with sufficiently high factor loadings, and the groupings of items on each domain aligned with what would be expected based on the content and themes of the questions [29]. This is an example of the structural validity of the stigma instrument. Moreover, the overall stigma scale and each subscale had acceptable alpha values, suggesting appropriate levels of internal consistency in these measures.

The current study results also exemplified convergent validity of the stigma measure in relation to expected variables. When collapsed across all three domains, overall stigma scores were positively associated with depression severity and negatively associated with education and adaptive coping behaviors. Based on the research literature, these associations were expected [17, 40]. In contrast, age was not clearly associated with overall stigma score, nor was sex, wealth, recent stressful life events, or social support. While the lack of strong or statistically significant associations with these constructs was not anticipated, the absence of such associations does make sense in context. First, given that the majority of study participants were female, there were fewer males in the study sample to compare stigma scores across these groups, leading to imprecise estimates of the association between sex and the outcome variables. Second, given that stigma is a multidimensional measure, relationships between these constructs may vary in direction across dimensions of the measure, ultimately averaging out to a null relationship between the construct and overall stigma score. One example of associations in opposite directions is the relationship between social support and disclosure carryover versus treatment carryover. In the disclosure carryover domain, individuals with stronger levels of social support, particularly from family, may be concerned about the impact of Thandi's disclosure on her family. Conversely, in the treatment carryover domain, which focused on the likelihood of Thandi losing friends due to receiving treatment, individuals with higher levels of social support, particularly from friends, may anticipate that the friends would

not abandon them for seeking help for their condition. The supposed mechanism behind these associations relies on the probable scenario that patients internalized Thandi's story and spoke based on their own current or anticipated experiences of depression-related stigma. Previous research supports that, when vignettes closely align with the experience of participants, their responses are related to their own lived experiences or their familiarity with others' experiences in these scenarios [41, 42]. Overall, these contradicting associations across domains, paired with the correlation matrix in Table 3, provide support for utilizing the domain-specific stigma scores in future multivariable analyses rather than using an overall stigma score. Our study results also exemplify the divergent validity of the measure, in that there was no association between stigma scores and height (adjusted for sex), just as would be expected.

Another important finding from this study was that most participants did not identify Thandi's condition as depression; rather, they used a variety of words to describe her state. While 37% of participants directly identified Thandi's condition as depression, another 32% described her condition as stress, and 7% called it the "normal ups and downs of life." These three concepts are not exclusive; stress and "ups and downs of life" may also be used to describe Thandi's situation and represent other commonly used terms to describe Thandi's condition within our patient population. A similar trend was observed in South Africa: in a survey where respondents were given vignettes with obvious presentations of depression and substance use disorder, they were more likely to attribute those symptoms to stress [17]. Further, in Zimbabwe, depression is commonly characterized as "thinking too much" [43]. These descriptions highlight the variation in language used around depressive symptoms amongst patients in Malawi, which is similar to other southern African countries. This variation in language used to describe depression underscores the necessity of incorporating patient screening for depression into standard care visits using validated tools in the patient population rather than assuming that patients who say they are just "stressed" are not struggling with depression. Moreover, this is an opportunity for patient education: patients who understand depression to be a medical condition that is amenable to treatment may be more inclined to seek or accept treatment as compared to those who believe such symptoms to be simply stress or an artifact of everyday life [28].

Limitations

There are several limitations to this study. First, our data were cross-sectional, and we are therefore unable to identify the temporal direction of the relationships between several of the constructs that were measured. Second, the SHARP cohort is a convenience sample, meaning that the sample

of patients enrolled in the SHARP study may differ from patients in other settings. As a result, this study may be susceptible to sampling bias, particularly because patients were being enrolled in a program related to depression, which may be a sensitive topic for some individuals. The SHARP study was also limited in the number of questions in the stigma tool due to an already extensive patient questionnaire within the broader study. Like any tool, this shortened stigma survey reflects the understanding and perspective of those who developed it and likely does not capture all aspects of stigma in Malawi. While this stigma tool does capture three dimensions of stigma in this population, a longer tool in future studies could provide deeper insight into the role of stigma in depression care.

As in all psychometric instruments, the validity of this stigma tool is also highly dependent on participants' interpretation of the prompts presented to them. In cross-cultural studies of validity, a further barrier is the translation of language and values from one context to another [44–46]. The stigma instrument was designed and discussed in English, then translated and edited by study team members fluent in the two target languages: Chichewa and Chitumbuka. The team members involved in the design of SHARP study materials chose to use a vignette format for this stigma instrument based on prior qualitative research that demonstrated that many people in Malawi may not have a clear understanding of what is meant when referring to “depression” [26]. The study team further elected to gender the vignette character as a woman for two primary reasons: First, traditional gender roles are very common in Malawi, and depression tends to impair function [47–52]; Second, the patient population was expected to include more women than men due to variation in healthcare-seeking behaviors by gender [53]. Thandi's gender inevitably influenced participants' interpretation of Thandi's depressive symptoms, as her struggle to complete the functions of her assigned gender role were deliberately described in the vignette. This likely influenced participants' responses to Thandi and the associated stigma prompts. However, we anticipate that such a bias in participant responses may be in the direction of less stigma, given that depressive symptoms tend to be more socially accepted in women as compared to men [54, 55]. Furthermore, the sex of the participant—which we used as a proxy for gender—did not have a significant relationship to participants' stigmatizing responses in this study (Table 4; Models 1–4), and the regression analysis results in Table 4 demonstrate that there was much imprecision in the coefficient associated with sex. If the gender of the participant strongly determined the way in which a participant responded to this vignette character, we would have expected a more precise estimate of the coefficient associated with sex in the regression analysis. Therefore, we do not anticipate that the gender of the vignette character differentially biased these analyses of the

reliability and validity of the stigma tool in our study sample. Nevertheless, in future iterations of this instrument, the study team will consider using a male vignette character to test the degree to which the vignette character's gender and traditional roles may influence participants' responses to stigma prompts.

While quantitative indicators are useful for understanding some aspects of the validity of psychometric instruments, some information is lost without qualitative data. Particularly in studies that adopt an instrument from one sociolinguistic context and then tailor that instrument to another, cognitive interviewing or other “thinking aloud” protocols are useful to understanding how participants are interpreting and responding to prompts [56, 57]. Therefore, the SHARP study team's next step in assessing this instrument's validity in this patient population is to analyze qualitative interviews among a subsample of SHARP patients and compare patients' qualitative responses to quantitative responses.

Strengths

Key strengths of this study include its sample size and breadth of data describing NCD patients with depressive symptoms in Malawi. Moreover, this study is the first to apply quantitative measurement of stigma in this patient population, offering further insight into depression-related stigma among individuals experiencing depressive symptoms, rather than describing stigma only in the general public.

Conclusion

The current study evaluates the psychometric properties of an instrument to measure depression-related stigma among individuals living with depression in Malawi. Validated tools are necessary to describe stigmatizing beliefs among stigmatized individuals and thereby better understand the unique barriers that these individuals face when accessing care or seeking support. Because stigma is multidimensional, we may not expect a patient with mostly treatment carryover concerns to have equal treatment adherence to a patient whose main concern is negative affect. It is therefore useful to have separate subscales of stigma for various dimensions, as exemplified in this analysis, to understand the ways in which these dimensions may differentially influence patient outcomes. Ultimately, longitudinal analyses of patient stigma and patient experiences throughout the treatment continuum are needed to provide critical information that may enhance the expansion of lifesaving psychiatric treatment in Malawi—validating a tool for measuring such stigma is a key first step.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00127-021-02202-w>.

Author contributions Study protocol and materials were developed by CZ, JM, HA, MU, MS, BNG, MH, BWP, and JM. Data analysis was performed by JD. The first draft of the manuscript was written by JD and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by the National Institute of Mental Health (U19MH113202-01; U19MH113202-03S1). JMD was also supported by funding from the National Institute of Allergy and Infectious Diseases (T32AI070114).

Availability of data and materials Data from this study are available via the National Institute of Mental Health Data Archive (NDA): https://nda.nih.gov/edit_collection.html?id=2822

Code availability Code will be made available upon request.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

Ethics approval This study has been approved by the University of North Carolina Biomedical Institutional Review Board (UNC IRB). It has also been approved by the Malawi National Health Science Research Committee (NHSRC). All study participants provided written informed consent and were provided with a small honorarium for participating in each research interview, as approved by the NHSRC and UNC IRB.

References

- Population, total—World | Data. In: The World Bank. <https://data.worldbank.org/indicator/SP.POP.TOTL?locations=1W>. Accessed 21 Sep 2020
- James SL, Abate D, Abate KH et al (2018) Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 392:1789–1858. [https://doi.org/10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7)
- Abas M, Bowers T, Manda E et al (2016) Opening up the mind: problem-solving therapy delivered by female lay health workers to improve access to evidence-based care for depression and other common mental disorders through the Friendship Bench Project in Zimbabwe. *Int J Mental Health Syst*. <https://doi.org/10.1186/s13033-016-0071-9>
- Chibanda D, Mesu P, Kajawu L et al (2011) Problem-solving therapy for depression and common mental disorders in Zimbabwe: piloting a task-shifting primary mental health care intervention in a population with a high prevalence of people living with HIV. *BMC Public Health* 11:1–10. <https://doi.org/10.1186/1471-2458-11-828>
- Goffman E (2009) *Stigma: Notes on the Management of Spoiled Identity*. Simon and Schuster
- Pescosolido BA, Martin JK, Long JS et al (2010) “A disease like any other”? A decade of change in public reactions to schizophrenia, depression, and alcohol dependence. *Am J Psychiatry* 167:1321–1330. <https://doi.org/10.1176/appi.ajp.2010.09121743>
- Barney LJ, Griffiths KM, Jorm AF, Christensen H (2006) Stigma about depression and its impact on help-seeking intentions. *Aust N Z J Psychiatry* 40:51–54. <https://doi.org/10.1080/j.1440-1614.2006.01741.x>
- Thorncroft G (2008) Stigma and discrimination limit access to mental health care. *Epidemiol Psychiatr Soc* 17:14–19
- Corrigan PW (1998) The impact of stigma on severe mental illness. *Cogn Behav Pract* 5:201–222. [https://doi.org/10.1016/S1077-7229\(98\)80006-0](https://doi.org/10.1016/S1077-7229(98)80006-0)
- Clement S, Schauman O, Graham T, et al What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. 17
- Sirey JA, Bruce ML, Alexopoulos GS et al (2001) Stigma as a barrier to recovery: perceived stigma and patient-rated severity of illness as predictors of antidepressant drug adherence. *PS* 52:1615–1620. <https://doi.org/10.1176/appi.ps.52.12.1615>
- Cavelti M, Kvrjic S, Beck E-M et al (2012) Self-stigma and its relationship with insight, demoralization, and clinical outcome among people with schizophrenia spectrum disorders. *Compr Psychiatry* 53:468–479. <https://doi.org/10.1016/j.comppsy.2011.08.001>
- Wirth JH, Bodenhausen GV (2009) The role of gender in mental-illness stigma: a national experiment. *Psychol Sci* 20:169–173. <https://doi.org/10.1111/j.1467-9280.2009.02282.x>
- Smith RA, Applegate A (2018) Mental health stigma and communication and their intersections with education. *Commun Educ* 67:382–393. <https://doi.org/10.1080/03634523.2018.1465988>
- Brown RL, Moloney ME, Brown J (2018) Gender differences in the processes linking public stigma and self-disclosure among college students with mental illness. *J Community Psychol* 46:202–212. <https://doi.org/10.1002/jcop.21933>
- Khan N, Kausar R, Khalid A, Farooq A (2015) Gender differences among discrimination & stigma experienced by depressive patients in Pakistan. *Pak J Med Sci* 31:1432–1436. <https://doi.org/10.12669/pjms.316.8454>
- Hugo CJ, Boshoff DEL, Traut A et al (2003) Community attitudes toward and knowledge of mental illness in South Africa. *Soc Psychiatry Psychiatr Epidemiol* 38:715–719. <https://doi.org/10.1007/s00127-003-0695-3>
- Crabb J, Stewart RC, Kokota D et al (2012) Attitudes towards mental illness in Malawi: a cross-sectional survey. *BMC Public Health* 12:541. <https://doi.org/10.1186/1471-2458-12-541>
- DeVellis RF (2017) *Scale development: theory and applications*, Fourth edition. SAGE, Los Angeles
- Udedi M, Muula AS, Stewart RC, Pence BW (2019) The validity of the patient health Questionnaire-9 to screen for depression in patients with type-2 diabetes mellitus in non-communicable diseases clinics in Malawi. *BMC Psychiatry*. <https://doi.org/10.1186/s12888-019-2062-2>
- Kroenke K, Spitzer RL, Williams JB (2001) The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 16:606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Levis B, Benedetti A, Thombs BD (2019) Accuracy of patient health questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. *BMJ*. <https://doi.org/10.1136/bmj.11476>
- Gilbody S, Richards D, Brealey S, Hewitt C (2007) Screening for depression in medical settings with the patient health questionnaire (PHQ): a diagnostic meta-analysis. *J Gen Intern Med* 22:1596–1602. <https://doi.org/10.1007/s11606-007-0333-y>
- Pescosolido BA, Martin JK, Olafsdottir S et al (2015) The theory of industrial society and cultural schemata: does the “Cultural

- Myth Of Stigma” underlie the WHO schizophrenia paradox? *AJS* 121:783–825. <https://doi.org/10.1086/683225>
25. Pescosolido ACK Bernice A (2020) Countries and Cultural Differences in the Stigma of Mental Illness: The East–West Divide—Anne C. Krendl, Bernice A. Pescosolido, 2020. *J Cross-Cultural Psychol*
 26. Stockton MA, Minnick CE, Kulisewa K et al (2021) A mixed-methods process evaluation: integrating depression treatment into HIV care in Malawi. *Global Health Sci Pract*. <https://doi.org/10.9745/GHSP-D-20-00607>
 27. Zimba CC, Akiba CF, Mawere M et al (2021) Facilitators, barriers and potential solutions to the integration of depression and non-communicable diseases (NCDs) care in Malawi: a qualitative study with service providers. *Int J Ment Health Syst* 15:59. <https://doi.org/10.1186/s13033-021-00480-0>
 28. Akiba CF, Zimba CC, Thom A et al (2020) The role of patient–provider communication: a qualitative study of patient attitudes regarding co-occurring depression and chronic diseases in Malawi. *BMC Psychiatry* 20:243. <https://doi.org/10.1186/s12888-020-02657-2>
 29. Pescosolido BA, Martin JK (2015) The stigma complex. *Annu Rev Sociol* 41:87–116. <https://doi.org/10.1146/annurev-soc-071312-145702>
 30. Griffiths KM, Christensen H, Jorm AF (2008) Predictors of depression stigma. *BMC Psychiatry* 8:25. <https://doi.org/10.1186/1471-244X-8-25>
 31. Shibre T, Negash A, Kullgren G et al (2001) Perception of stigma among family members of individuals with schizophrenia and major affective disorders in rural Ethiopia. *Soc Psychiatry Psychiatr Epidemiol* 36:299–303. <https://doi.org/10.1007/s001270170048>
 32. Leserman J, Ironson G, O’Cleirigh C et al (2008) Stressful life events and adherence in HIV. *AIDS Patient Care STDS* 22:403–411. <https://doi.org/10.1089/apc.2007.0175>
 33. Sarason IG, Johnson JH, Siegel JM (1979) Assessing the impact of life changes: development of the life experiences survey. *J Consult Clin Psychol* 46:932. <https://doi.org/10.1037/0022-006X.46.5.932>
 34. Stewart RC, Umar E, Tomenson B, Creed F (2014) Validation of the multi-dimensional scale of perceived social support (MSPSS) and the relationship between social support, intimate partner violence and antenatal depression in Malawi. *BMC Psychiatry* 14:180. <https://doi.org/10.1186/1471-244X-14-180>
 35. Zimet GD, Dahlem NW, Zimet SG, Farley GK (1988) The multidimensional scale of perceived social support. *J Pers Assess* 52:30–41. https://doi.org/10.1207/s15327752jpa5201_2
 36. Carver CS (1997) You want to measure coping but your protocol’ too long: consider the brief cope. *Int J Behav Med* 4:92. https://doi.org/10.1207/s15327558ijbm0401_6
 37. Cattell RB (1966) The scree test for the number of factors. *Multivar Behav Res* 1:245–276. https://doi.org/10.1207/s15327906mbr0102_10
 38. Knekta E, Runyon C, Eddy S (2019) One size doesn’t fit all: using factor analysis to gather validity evidence when using surveys in your research. *LSE* 18:rm1. <https://doi.org/10.1187/cbe.18-04-0064>
 39. (2019) Stata Statistical Software: Release 16. StataCorp, LLC, College Station
 40. Alonso J, Buron A, Rojas-Farreras S et al (2009) Perceived stigma among individuals with common mental disorders. *J Affect Disord* 118:180–186. <https://doi.org/10.1016/j.jad.2009.02.006>
 41. Rahman N (1996) Caregivers’ sensitivity to conflict: the use of the vignette methodology. *J Elder Abuse Negl* 8:35–47. https://doi.org/10.1300/J084v08n01_02
 42. Crafter S, de Abreu G, Cline T, O’Dell L (2015) Using vignette methodology as a tool for exploring cultural identity positions of language brokers. *J Constr Psychol* 28:83–96. <https://doi.org/10.1080/10720537.2014.923354>
 43. Patel V, Abas M, Broadhead J et al (2001) Depression in developing countries: lessons from Zimbabwe. *BMJ* 322:482–484
 44. King KM, Khan N, LeBlanc P, Quan H (2011) Examining and establishing translational and conceptual equivalence of survey questionnaires for a multi-ethnic, multi-language study. *J Adv Nurs* 67:2267–2274. <https://doi.org/10.1111/j.1365-2648.2011.05679.x>
 45. (2000) A construct-based approach to equivalence: methodologies for cross-cultural/multicultural personality assessment research. In: *Handbook of Cross-Cultural and Multicultural Personality Assessment*. Routledge
 46. Laher S, Cockcroft K (2013) Current and future trends in psychological assessment in South Africa: challenges and opportunities. In: Laher S, Cockcroft K (eds) *psychological assessment in South Africa*. Wits University Press, pp 535–552
 47. Manda-Taylor L, Mwale D, Phiri T et al (2017) Changing times? Gender roles and relationships in maternal, newborn and child health in Malawi. *BMC Pregnancy Childbirth* 17:321. <https://doi.org/10.1186/s12884-017-1523-1>
 48. Ng’oma M, Meltzer-Brody S, Chirwa E, Stewart RC (2019) “Passing through difficult times”: perceptions of perinatal depression and treatment needs in Malawi—a qualitative study to inform the development of a culturally sensitive intervention. *PLoS ONE* 14:e0217102. <https://doi.org/10.1371/journal.pone.0217102>
 49. Kumwenda MK, Corbett EL, Chikovore J et al (2018) Discordance, disclosure and normative gender roles: barriers to couple testing within a community-level HIV self-testing intervention in Urban Blantyre, Malawi. *AIDS Behav* 22:2491–2499. <https://doi.org/10.1007/s10461-018-2038-0>
 50. Ulbricht CM, Rothschild AJ, Lapane KL (2016) Functional impairment and changes in depression subtypes for women in STAR*D: a latent transition analysis. *J Womens Health (Larchmt)* 25:464–472. <https://doi.org/10.1089/jwh.2015.5361>
 51. Habtamu K, Medhin G, Selamu M et al (2019) Functional impairment among people diagnosed with depression in primary healthcare in rural Ethiopia: a comparative cross-sectional study. *Int J Ment Health Syst* 13:1–12. <https://doi.org/10.1186/s13033-019-0305-8>
 52. Hammer-Helmich L, Haro JM, Jönsson B et al (2018) Functional impairment in patients with major depressive disorder: the 2-year PERFORM study. *Neuropsychiatr Dis Treat* 14:239–249. <https://doi.org/10.2147/NDT.S146098>
 53. Yeatman S, Chamberlin S, Dovel K (2018) Women’s (health) work: a population-based, cross-sectional study of gender differences in time spent seeking health care in Malawi. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0209586>
 54. Townsend L, Musci R, Stuart E et al (2019) Gender differences in depression literacy and stigma following a randomized controlled evaluation of a universal depression education program. *J Adolesc Health* 64:472–477. <https://doi.org/10.1016/j.jadohealth.2018.10.298>
 55. Sorsdahl K, Stein DJ, Myers B (2012) Negative attributions towards people with substance use disorders in South Africa: variation across substances and by gender. *BMC Psychiatry* 12:1–8. <https://doi.org/10.1186/1471-244X-12-101>
 56. Güss CD (2018) What is going through your mind? Thinking aloud as a method in cross-cultural psychology. *Front Psychol* 9:1292. <https://doi.org/10.3389/fpsyg.2018.01292>
 57. Willis G (2005) *Cognitive Interviewing*. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States of America