

# Developing a Global Cancer Stigma Index

SAGE Open  
July-September 2014: 1–9  
© The Author(s) 2014  
DOI: 10.1177/2158244014547875  
sgo.sagepub.com



**Maria Orlando Edelen<sup>1</sup>, Anita Chandra<sup>2</sup>, Brian Stucky<sup>3</sup>,  
Rebekkah Schear<sup>4</sup>, Claire Neal<sup>4</sup>, and Ruth Rechis<sup>4</sup>**

## Abstract

Despite increasing recognition about the stigma associated with cancer screening, diagnosis, and treatment-seeking behaviors, there has been relatively little attention paid to how to assess and intervene to reduce that stigma. An index to measure cancer stigma could empower health program developers and policymakers by identifying the key areas in which a population could benefit from education to change perceptions and address misinformation. The index also could be used to rank countries and communities based on their level of cancer stigma to assess where interventions are needed. We used structured literature review and expert review to generate a cancer stigma item pool. The item pool was subject to cognitive interviews for cultural appropriateness and comprehension; and data from initial pilot testing were used to reduce the pool of items for translation and field testing. The field test was conducted using a web-based survey in four samples representing two regions and three languages—English and Arabic speakers in Jordan and Egypt, and English and Mandarin Chinese speakers in China. Factor analyses and item response theory were applied to finalize the index. The analyses resulted in a 12-item cancer stigma index (CSI) that was reliable across all four samples. The CSI scores were highly correlated with a general illness stigma scale, and operated as expected noting higher cancer stigma among men and those with lower income. The CSI can be used to inform initial cancer education efforts, identifying overall stigma levels in a country or community and particular issue areas requiring intervention.

## Keywords

cancer, stigma, health, chronic disease, help-seeking, health disparities

## Introduction

Over the last several years, stigma associated with health-seeking behaviors has received increasing attention. Health-related stigma (or principally, disease-related stigma) is unique and presents significant challenges and barriers for the global health community to overcome. The framing of health-related stigma has begun to advance a more complex discussion of stigma, one that encompasses both the internalization of stigma by the individual and the public reaction and potential marginalization that may occur. First, the word “disease” alone can induce a sense of stigma (Green, Davis, Karshmer, Marsh, & Straight, 2005; Pettit, 2008). There are a number of diseases that have historically been highly stigmatized, including mental health disorders, HIV/AIDS, sexually transmitted diseases, leprosy, and skin diseases (Greene & Banerjee, 2006; Sartorius, 2007). An emerging body of data in the stigma literature indicates that cancer is also often among the diseases that is highly stigmatized (LIVESTRONG Foundation, 2007), yet it is less researched than other health issues. Individuals often react to that stigma by making decisions about whether or not to disclose their condition or seek treatment (Joachim & Acorn, 2000).

Specific types of cancer may also carry disease-specific stigma. For example, cervical and lung cancer are often cited because each is linked to behavior that may be deemed undesirable or marginal. In the case of lung cancer, individuals may feel guilt and shame attributed to their diagnosis, due to the link between smoking and cancer. Guilt may lead to denial of the diagnosis until such a point that treatment may not be successful (Batson et al., 1997). Cervical cancer, breast cancer, and uterine cancer may also carry a particular stigma as these cancers are often linked to sexual health, regardless of the actual disease pathway. In some patriarchal societies, women are considered to be the property of their spouses, and must comply with their spouse's wishes to not seek treatment. Religious and cultural beliefs may prohibit

<sup>1</sup>RAND Corporation, Boston, MA, USA

<sup>2</sup>RAND Corporation, Arlington, VA, USA

<sup>3</sup>RAND Corporation, Santa Monica, CA, USA

<sup>4</sup>Livestrong Foundation, Austin, TX, USA

### Corresponding Author:

Maria Orlando Edelen, RAND Corporation, 20 Park Plaza, Suite 920,  
Boston, MA 02116, USA.

Email: [Maria\\_Edelen@rand.org](mailto:Maria_Edelen@rand.org)



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 3.0 License

(<http://www.creativecommons.org/licenses/by/3.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<http://www.uk.sagepub.com/aboutus/openaccess.htm>).

seeking medical attention for parts of the body having a sexual connotation (Brewster & Moradi, 2010).

Despite this emerging recognition that stigma related to cancer can deter critical health-seeking behaviors, there has been comparatively little effort to measure the level of cancer stigma in a given population or community to intervene. Most stigma scales have focused in other areas, principally HIV/AIDS and mental health (Brohan, Slade, Clement, & Thornicroft, 2010; Evans-Lacko et al., 2010; Uys et al., 2009). This study sought to develop a cancer stigma index (CSI) to measure perspectives on cancer, specifically attitudes about cancer screening and treatment, and to help inform awareness and education programs. We proceeded in two phases. First, we conducted an initial study to gather stigma measures and test a preliminary item pool (Study 1). Then, we conducted a full field test in two regions (Study 2).

## Study 1: Creation of a Cancer Stigma Item Pool and Initial Pilot Test

We used established methods for item pool development following the National Institutes of Health Patient-Reported Outcomes Measurement Information System (NIH PROMIS®) initiative blueprint (Reeve et al., 2007). The main goals of PROMIS® (<http://www.nihpromis.org/default.aspx>), which is part of the NIH's Roadmap initiative, are to standardize a set of assessment tools and to use item response theory (IRT) techniques and advances in computer technology to create brief yet highly reliable and flexible assessment tools to measure patient-reported outcomes (Ader, 2007; Cella et al., 2007; Fries, Bruce, & Cella, 2005). Due to its scope and success, the rigorous approach utilized by PROMIS has become something of a standard for modern instrument development (DeWalt, Rothrock, Yount, & Stone, 2007).

Briefly, we first conducted a structured literature search for stigma as it relates to cancer, HIV/AIDS, mental health, and other health issues. The initial search summarized literature in these areas known to underlie stigma: information and myths, fear, shame and labeling, concerns about diagnosis, concerns about treatment-seeking, concerns about peer and family disclosure, and concerns about public disclosure. We identified and organized a total of 553 items from 29 measures into four broad domains: characterization about those with the disease, self-stigma among those with the disease, expectations of what others may think of those with the disease, and positive views on those with the disease. After review by five technical advisors, we reduced the number of items based on item scope and quality of item wording, and reworded them to be appropriate for the cancer stigma context and in comparable format.

### Cognitive Interviews

Using the reduced set of items, we conducted seven cognitive interviews to (a) assess whether the items were

comprehensible, (b) understand how respondents interpreted the items, and (c) ensure that the item content and wording were culturally appropriate. Because our initial pilot test was to be conducted in the Middle East, cognitive interview participants, although U.S.-based, were of Arab heritage, had been raised in the region (at least up to age 18), considered themselves culturally Arab, and spoke both English and Arabic. All respondents received a US\$25 gift card for their time, and all procedures were approved by the RAND Institutional Review Board (IRB). Following review of the cognitive interview transcripts, we made changes to the item pool based on respondent feedback including dropping items that were considered redundant and generating items to reflect stigma and religion, a theme that was considered important by interviewees.

### Pilot Test

#### Method

**Measures.** *Cancer stigma item pool.* A set of 59 candidate items, formatted using Likert-type response options (1 = *not at all* to 5 = *very much*), were administered to all respondents.

**General illness stigma.** As there were no items available on general illness stigma that corresponded well to the cancer stigma constructs we are trying to measure, we used information in the literature to develop five items reflecting attitudes of stigma toward general illness (e.g., If I had an illness, I would feel left out of things). These items also used the five-category Likert-type response format.

**Demographics.** All respondents provided basic demographic information, including gender, age, education, religion, country of origin, country of residence, and language spoken at home. Six additional items indicated the respondents' connections to cancer (e.g., I know someone with cancer, I am a caregiver of someone with cancer).

### Sample and Procedures

All study procedures were approved by the RAND IRB. We contracted with Harris Interactive to pilot the items via the Internet to English-speaking adult respondents residing in Egypt and Jordan. We collected a total of  $N = 1,016$  completed web surveys. The majority of respondents were male (72%), and the sample was also skewed toward younger and more educated individuals than is representative of the general population in the region (52% were <30 years old, 94% were <50 years old; 71% had university education). We addressed this imbalance by selecting a subset of the data that had the same number of men and women (294 each; Total  $N = 588$ ). All women from the original sample were retained, and we used a random sampling approach to select a subset of 294 males that was stratified according

to education and an indicator of having lived outside the country. Despite this stratification, the analytic sample was still relatively young (49% below 30 years of age) and fairly well-educated (68% completed some post-secondary education). Nineteen percent of respondents were from Jordan, 60% lived in urban areas, 95% were Muslim, 22% had lived outside the country for more than 5 years, and 59% reported some personal cancer connection.

### *Evaluation and Reduction of Item Pool*

For analyses, we randomly split the pilot sample of  $N = 588$  into two analytic samples for exploratory ( $n = 400$ ) and confirmatory ( $n = 188$ ) analyses. Using the exploratory sample, we conducted exploratory factor analysis (EFA) with Mplus software (Muthén & Muthén, 2007) modeling the 59 stigma items as categorical with the weighted least squares means and variance adjusted [WLSMV] estimator. The main goals of this analysis were to identify the structure of the item set and remove items that were not performing well. Following this item reduction, we conducted a confirmatory factor analysis (CFA) using the confirmatory sample ( $n = 188$ ), and evaluated model fit with standard diagnostic fit indices (root mean square error of approximation [RMSEA]  $\leq .08$ , Tucker–Lewis Index [TLI]  $\geq .95$ , comparative fit index [CFI]  $\geq .95$ ; Browne, Cudeck, Bollen, & Long, 1993; Hu & Bentler, 1999).

Further item reduction was achieved based on consideration of the overall goals of the index. In addition to examination of CFA results, we examined results from a series of IRT calibrations (conducted using IRTPRO; Cai, du Toit, & Thissen, 2011), including item properties, item fit, and local dependence indices to identify redundant or poorly performing items either in content or in terms of item properties. After discussion of results among study team members, the refined index was finalized for the next phase of field testing.

### *Factor Analyses*

Results from the initial EFA of the 59 items indicated that a two-factor solution was most appropriate. In this solution, the first factor consisted of 36 items reflecting negative stigma (fear, lack of understanding, negativity) and the second factor contained 23 items reflecting more positive statements (compassion for cancer, understanding, pragmatism). At this stage, we elected to remove a total of 7 items that either did not load cleanly on a single factor (double-loaders, 3 items) or loaded weakly on their respective factors (4 items).

The 52 remaining items were subject to a two-factor CFA. The fit of the initial model was not quite acceptable (CFA = .874, TLI = .869, RMSEA = .063). Model fit diagnostics suggested removal of two items from Factor 2, and this suggestion was supported by the fact that these were the only reverse-keyed items in that factor. After their removal, the fit

of the two-factor model was adequate (CFA = .905, TLI = .901, RMSEA = .057). The final solution had 31 items in Factor 1 and 19 items in Factor 2. After consultation, we elected to set aside all the Factor 2 items as our ultimate goal was to produce a unidimensional index, and the items in Factor 1 reflected the content of primary interest.

### *IRT Analyses*

Results from a series of IRT calibrations of the 31 remaining items from Factor 1 led to removal of 6 additional items based on poor fit to the IRT model and excess local dependence. After examining results from an IRT calibration of the remaining 25 items, we elected to retain all of these items for the larger field test. However, 3 were reworded for clarification and a new item was created to represent conflating cancer with death, as this was of particular interest. Finally, although we wished to arrive at a final index measuring only a single dimension of cancer stigma, we elected to reword and retain 5 items from the original Factor 2 that had desirable content; all other items from Factor 2 were discarded.

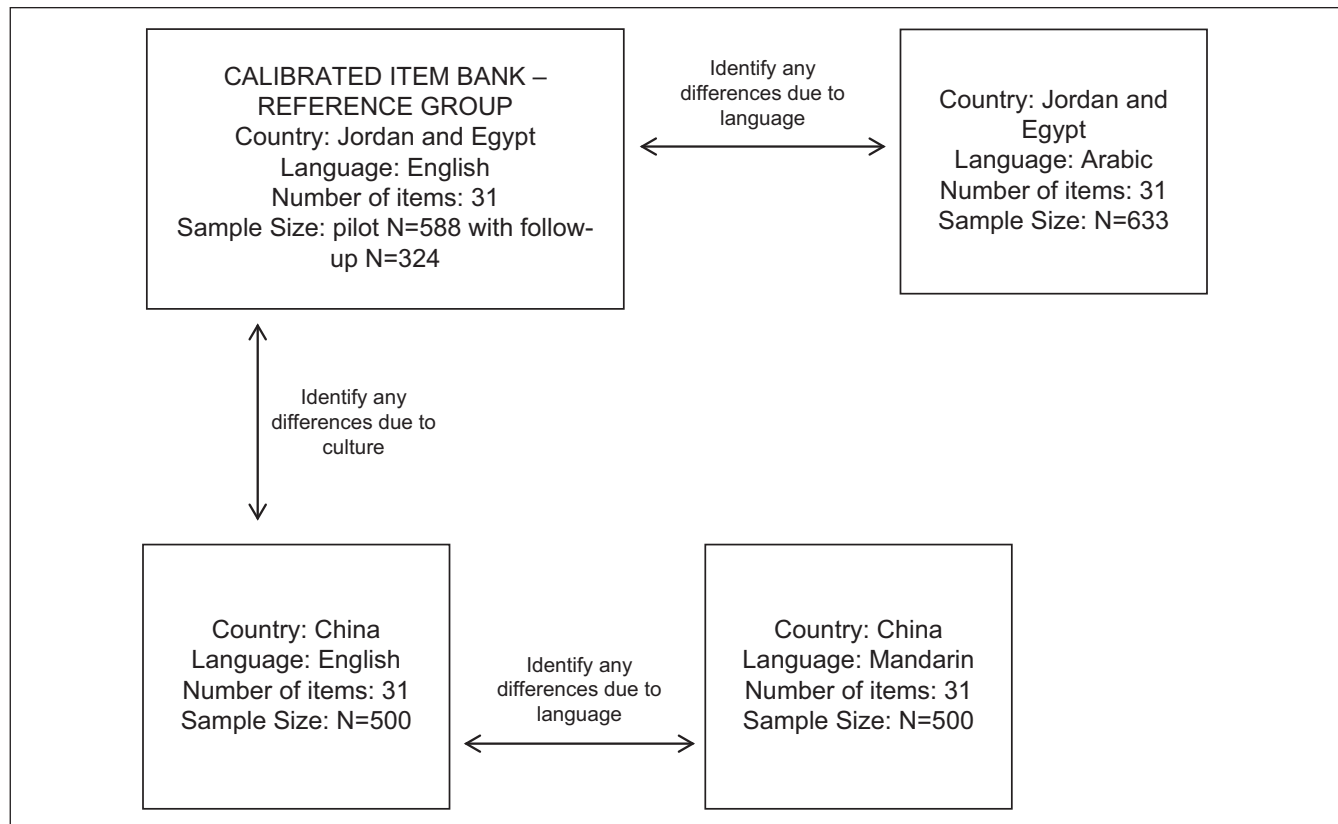
## **Study 2: Translation, Review, and Field Testing in Jordan/Egypt and China**

Based on strategic priorities, the revised 31-item index was prepared to be fielded among English- and Arabic-speaking respondents in the Middle East (Jordan/Egypt), and among English- and Mandarin-speaking respondents in China. Thus, the item set was translated into Arabic and simple Mandarin Chinese, and cognitive interviews were conducted for each translation of the instrument.

### *Cognitive Interviews*

*English- and Arabic-speaking, Arab respondents.* We conducted interviews with eight respondents of Arab heritage, who spoke both English and Arabic. All respondents reviewed the English version, and six reviewed both the English and Arabic versions. We mirrored the recruitment strategy used in Study 1. Respondents identified a few instances where the Arabic translation did not adequately capture the original content; these translated items were modified.

*English- and Mandarin Chinese-speaking, Chinese respondents.* We utilized professional, family, and peer networks to identify a demographically diverse set of individuals to participate in the cognitive interviews to review English and Mandarin Chinese versions of the cancer stigma item set. We conducted interviews via Skype which allowed us to include eight respondents in China as well as three in the United States. The sample included seven men and four women who were fairly well-distributed by age. In the Chinese interviews, respondents expressed confusion about items regarding



**Figure 1.** Flowchart of cancer stigma index development design.

isolation and being an outcast. Thus, we modified the translation to more clearly communicate these terms.

### Field Test

#### Method

**Measures.** All Study 2 measures were identical to Study 1 measures with the exception that the reduced 31 cancer stigma items were administered as opposed to the initial 59 items.

**Sample and procedures.** As can be seen in Figure 1, the 31-item field test was administered in two regions, Jordan/Egypt and China, and in three languages, English, Arabic, and Mandarin, to produce four distinct samples. Respondents residing in Jordan/Egypt were administered an English (JE,  $n = 324$ ) or an Arabic version (JA,  $n = 633$ ); and respondents residing in China were administered an English (CE,  $n = 500$ ) or a Mandarin version (CM,  $n = 500$ ). As in Study 1, we contracted with Harris Interactive to administer the field test via the Internet in both regions, and all study procedures were approved by the RAND IRB. The characteristics of the four field test samples are displayed in Table 1.

**Evaluation of item set.** Our first analytic step was to conduct CFAs with each sample to evaluate the extent to which

the 31 items represent a single dimension. Based on results from these analyses, we considered items for removal to improve unidimensionality across the four samples.

Once a set of items was identified that appeared to be sufficiently unidimensional in all samples, we used IRTPRO to conduct differential item functioning (DIF) analyses within an IRT framework. DIF, also referred to as measurement bias, occurs when people from different groups (e.g., gender or ethnicity) with the same level of the latent trait (in this case cancer stigma) have a different probability of giving a certain response to an item. Thus, we compared performance of (a) the Arabic and English items from the two Jordan/Egypt samples (JE–JA), (b) the Mandarin and English items from the two China samples (CE–CM), and (c) the English items from Jordan/Egypt and China samples (JE–CE) to determine the comparability of cancer stigma items across language and region (this process is depicted in Figure 1).

DIF analysis used three steps. First, two-group chi-square tests from IRTPRO were evaluated across comparison groups and the significance tests for all comparisons were adjusted using the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995) at  $p < .05$  to identify candidate items for removal. Next, to evaluate the magnitude of DIF, items demonstrating significant DIF after  $p$  value correction were further evaluated by computing the weighted



**Table 1.** Characteristics of Study 2 Field Test Sample, by Region and Language.

Variable	Value	Jordan/Egypt: English (n = 324)	Jordan/Egypt: Arabic (n = 633)	China: English (n = 500)	China: Mandarin (n = 500)
Gender	Male	224 (69.1%)	351 (55.5%)	251 (50.2%)	261 (52.2%)
	Female	100 (30.9%)	282 (44.6%)	249 (49.8%)	239 (47.8%)
Age	18-29	150 (46.3%)	244 (38.6%)	146 (29.2%)	136 (27.2%)
	30-49	147 (45.3%)	331 (52.3%)	289 (57.8%)	264 (52.8%)
	50+	27 (8.3%)	58 (9.2%)	65 (13.0%)	100 (20.0%)
Urbanicity	City/large town	251 (77.5%)	453 (71.6%)	404 (80.8%)	402 (80.4%)
	Suburb	28 (8.6%)	88 (13.9%)	53 (10.6%)	47 (9.4%)
	Small town	31 (9.6%)	25 (3.9%)	33 (6.6%)	37 (7.4%)
	Village/rural	14 (4.3%)	67 (10.6%)	10 (2.0%)	14 (2.8%)
Low income ( $\leq 800$ JOD, $\leq 3,000$ EGP, $\leq ¥84,999$ )	Yes	80 (24.7%)	200 (31.6%)	197 (39.4%)	184 (36.8%)
	No	244 (75.3%)	433 (68.4%)	303 (60.6%)	316 (63.2%)
Post-secondary education	Yes	295 (91.6%)	547 (87.9%)	429 (87.9%)	424 (85.7%)
	No	27 (8.4%)	85 (13.5%)	59 (12.1%)	71 (14.3%)
Religion	Islam	279 (86.1%)	577 (91.2%)	10 (2.0%)	3 (.6%)
	Buddhism	0 (0%)	0 (0%)	54 (10.8%)	89 (17.8%)
	Christianity	34 (10.5%)	27 (4.3%)	48 (9.6%)	26 (5.2%)
	Other/multiple/ decline	11 (3.4%)	29 (4.6%)	122 (24.4%)	47 (9.4%)
	None	0 (0%)	0 (0%)	266 (53.2%)	335 (67.0%)

area between the expected score curves (wABC; Edelen, Stucky, & Chandra, 2013).

After removing items with problematic DIF, we fit a four-group IRT model. Based on those results, we reconsidered item content and properties in an effort to further reduce the item set and arrive at a final index.

**Scoring and examining CSI scores.** Once the CSI was finalized, we calculated IRT-based scores (i.e., expected a posteriori [EAP]) using a summed score conversion algorithm (Thissen, Pommerich, Billeaud, & Williams, 1995). These scores retain the benefits of the IRT model and are also practical for general use via score translation tables. Summed score EAPs were generated for the CSI using the final four-group IRT model and were rescaled along a *T*-score metric with the Jordan English sample mean set to 50 and standard deviation to 10.

Finally, we conducted a set of descriptive analyses to examine the CSI scores according to demographic and personal characteristics of interest (e.g., gender, income, age, personal experience with cancer, attitudes about general health stigma).

## Results

**Identifying the 12-item CSI.** The fit of the 31 items to a single-factor model in each sample was reasonable but did not reach standard criteria for all four samples (JE: CFI = .886, TLI = .878, RMSEA = .090; JA: CFI = .852, TLI = .841, RMSEA = .096; CE: CFI = .885, TLI = .877, RMSEA = .084; CM: CFI = .921, TLI = .916, RMSEA = .083), indicating the need for some modifications. In consultation with the larger study team, 4 items were removed to obtain acceptable

**Table 2.** Items Comprising the CSI.

1. I would feel uncomfortable talking to a person with cancer
2. Treatment and support are useless for someone with cancer
3. I would feel uncomfortable sitting next to someone with cancer
4. I would feel uncomfortable sending own child to school with another child with cancer
5. I would feel uncomfortable if someone with cancer lived nearby
6. If a close friend had cancer, I would avoid him or her
7. I would feel uncomfortable being friends with someone with cancer
8. People can only blame themselves for getting cancer
9. If had cancer, I would be ashamed
10. I would feel isolated/alone if I received treatment for cancer
11. If my spouse had cancer, I would be ashamed of him or her
12. If my spouse had cancer, I would consider leaving him or her

Note. CSI = cancer stigma index.

model fit with the same set of items across all four samples (JE: CFI = .955, TLI = .951, RMSEA = .063; JA: CFI = .922, TLI = .916, RMSEA = .078; CE: CFI = .903, TLI = .895, RMSEA = .087; CM: CFI = .946, TLI = .941, RMSEA = .078). IRT-DIF analyses identified 9 items with problematic DIF that were also removed.

We used results from a four-group IRT calibration of the remaining 18 items to reduce the item set further by removing items based on redundant content and/or poor psychometric properties. In all, we removed 6 items at this stage. Items comprising the final 12-item CSI are listed in Table 2.

**Table 3.** Cancer Stigma Scores Among Demographic Groups for Each Sample and Combined.

	Combined Sample	Jordan/Egypt: English	Jordan/Egypt: Arabic	China: English	China: Mandarin
<b>Gender</b>					
$M_M (SD_M)$	52.0 (10.3)	50.3 (9.9)	48.9 (7.5)	58.4 (12.0)	50.7 (8.6)
$M_F (SD_F)$	50.4 (9.4)	49.2 (10.3)	46.5 (6.1)	57.6 (10.6)	49.2 (7.4)
$p$	.0004	.35	<.0001	.46	.03
<b>Income</b>					
$M_H (SD_H)$	50.7 (9.7)	48.6 (8.8)	46.9 (6.0)	58.5 (11.7)	50.0 (8.0)
$M_L (SD_L)$	52.4 (10.2)	54.2 (12.3)	49.1 (8.2)	57.3 (10.7)	49.9 (8.4)
$p$	.0003	.0003	.0007	.27	.88
<b>Age</b>					
$M_L (SD_L)$	51.0 (9.5)	52.1 (11.2)	48.1 (7.1)	56.0 (9.9)	49.7 (8.1)
$M_M (SD_M)$	51.6 (10.4)	48.6 (8.9)	47.4 (6.9)	59.1 (11.9)	50.3 (8.4)
$M_H (SD_H)$	50.4 (9.0)	45.9 (5.4)	46.1 (5.0)	57.7 (11.1)	49.4 (7.1)
$p$	.18	.0009	.10	.03	.60
<b>Post-secondary education</b>					
$M_N (SD_N)$	51.9 (11.3)	59.0 (15.4)	46.7 (7.1)	59.3 (12.1)	49.4 (8.3)
$M_Y (SD_Y)$	51.1 (9.5)	49.2 (9.0)	47.7 (6.8)	57.6 (11.0)	50.1 (8.0)
$p$	.27	.003	.21	.28	.52
<b>Lived outside country &gt;5 years</b>					
$M_N (SD_N)$	50.1 (8.8)	49.9 (10.0)	47.6 (6.5)	55.7 (10.3)	49.2 (7.5)
$M_Y (SD_Y)$	53.7 (11.5)	50.1 (10.1)	47.6 (8.1)	60.9 (12.0)	51.6 (9.0)
$p$	<.0001	.85	.97	<.0001	.005
<b>Diagnosed with cancer: Self</b>					
$M_N (SD_N)$	50.6 (9.2)	49.6 (9.8)	47.5 (6.6)	56.7 (10.5)	49.9 (7.8)
$M_Y (SD_Y)$	64.4 (14.3)	57.1 (12.6)	52.1 (18.5) <sup>a</sup>	67.4 (12.6)	67.4 (24.4) <sup>a</sup>
$p$	<.0001	.005	.54	<.0001	.34
<b>Diagnosed with cancer: Family/loved one</b>					
$M_N (SD_N)$	52.1 (10.3)	51.6 (11.0)	48.4 (7.7)	58.1 (11.4)	50.1 (8.3)
$M_Y (SD_Y)$	49.4 (8.6)	47.4 (7.5)	46.4 (5.2)	57.6 (11.1)	49.6 (7.5)
$p$	<.0001	<.0001	.0002	.63	.56
<b>Ever paid caregiver for cancer patient</b>					
$M_N (SD_N)$	50.2 (8.9)	48.9 (8.9)	47.5 (6.8)	56.1 (10.6)	49.6 (7.6)
$M_Y (SD_Y)$	55.2 (12.2)	51.8 (11.2)	47.9 (7.6)	62.1 (11.8)	52.0 (10.7)
$p$	<.0001	.05	.64	<.0001	.09

<sup>a</sup>Group size < 10.

We ran a final four-group calibration to obtain item parameters and generate CSI scores. Appendix A1 provides a score translation table for the 12-item CSI. The IRT-based score reliability for the CSI varies by sample and ranges from acceptable to excellent (JE = .79; JA = .73; CE = .91; CM = .81).

**Examining CSI scores.** The JA sample had the lowest CSI score ( $M = 47.6$ ,  $SD = 6.9$ ), with JE ( $M = 50.0$ ,  $SD = 10.0$ ) and CM ( $M = 50.0$ ,  $SD = 8.1$ ) both at 50. The CE sample had the highest mean CSI score ( $M = 58.0$ ,  $SD = 11.3$ ). Tests of significance between the four CSI sample means revealed that the JE and CM means are not different from one another, but all other group mean comparisons are statistically significant.

Correlation analyses revealed a moderate to strong correlation between the CSI and general illness stigma across all

four samples, providing some preliminary validity evidence for the CSI (range across samples  $r = .35-.50$ ). The correlation of the CSI with the item likening cancer to a death sentence was less consistent and slightly lower on average (range across samples  $r = .20-.49$ ). All correlations were significantly different from 0 at  $p < .05$ .

Scores from the CSI were compared across various demographic groups for each of the four samples to establish initial validity evidence. A summary of these results is contained in Table 3.

## Discussion

This analysis of cancer stigma and development of the CSI will provide critical benefits to the cancer research and control fields. Our process of integrating literature review with stakeholder input and measures analysis represents a robust

method of developing a quality, user-friendly index that can be used by cancer organizations to inform cancer stigma reduction initiatives and broader public awareness campaigns. Moreover, the CSI can be added to cancer research studies examining patient, family, and public perspectives regarding cancer screening, diagnoses, and treatment.

While there were limited cancer stigma items or scales available for modification at the outset of this study, there were several scales in the area of stigma and health care decision making as well as stigma and chronic health issues (e.g., mental health) that were particularly useful. The stakeholder input we solicited through technical advisors and cognitive interviews was critical. Without that feedback, the scale would not have included particular cultural “pulse points” such as the role of religion or fate in driving cancer views. The moderate to strong correlation between the CSI and general illness stigma across all four samples suggests that the CSI is reasonably robust and indicative of general health-seeking stigma. As expected, male gender and low income were associated with higher CSI scores. Interestingly, CSI scores indicated that those who had a loved one diagnosed with cancer reported lower stigma, whereas those who had personally experienced cancer reported more stigma.

Our approach attempted to limit potential weaknesses in design where possible. But a few study limitations should be noted. First, our cognitive interviews, particularly for the Arab origin samples, were not conducted in Egypt and Jordan. Although we attempted to find individuals who had strong cultural ties (e.g., using criteria about upbringing in the region), we may not have received the full complement of cultural insights from those who were Americans or had spent considerable time in the United States. Second, our pilot and field tests endeavored to obtain diversity by age, gender, income, and education. For the latter two categories, we approached our goal but did not always meet it in terms of education level, with a slightly higher education status overall. We know that education may influence cancer stigma perspectives. Furthermore, our mode of testing the cancer stigma items was web-based, which may impede participation from those of lower socioeconomic status. Mode effects of web-based administration could not be tested (given that was the only mode used), but may affect the interpretation of items.

More robust validity tests of the CSI will require additional use in the countries in which we developed the first versions of the scale—Egypt, Jordan, and China. This testing may include using the CSI in diverse communities and with a wide variety of subpopulations (e.g., setting, age). The scale was developed for use initially in these countries based on strategic plans and investments. However, the intention is for the scale to ultimately be used worldwide. Thus, as the CSI is translated and used in new contexts, it will be important to step through all of the phases used in this study, principally review of the translation by experts and some version

of cognitive interviews to ensure terms and whole items are interpreted as intended. Furthermore, data from field tests in new regions and languages must be analyzed to determine the comparability of CSI scores back to the reference sample (Jordan/Egypt—English).

It is important and possible to show that CSI scores correlate with other indicators of stigma. For example, we should expect to find that higher CSI scores correlate with national or local policies that discriminate against people with cancer, and predict lower levels of treatment-seeking, less positive psychological well-being, and greater social isolation among people with cancer. Future research along these lines could provide valuable validity evidence for the CSI. Overall, the CSI can be used to inform initial cancer education efforts, identifying overall stigma levels in a country or community and particular issue areas requiring concerted intervention. Over time, following careful data analyses and perhaps slight modifications to index scoring, the CSI can be used as an index comparing countries or communities on stigma levels, prioritizing where education resources should be allocated, and helping to determine the impact of stigma reduction efforts.

## Appendix A1

### *12-Item Cancer Stigma Index Total Score to T-Score Translation Table*

Total Score	T-Score	Total Score	T-Score
12	42	37	69
13	43	38	70
14	44	39	71
15	46	40	72
16	47	41	73
17	48	42	74
18	49	43	75
19	50	44	76
20	51	45	77
21	52	46	78
22	53	47	80
23	54	48	81
24	55	49	82
25	56	50	83
26	57	51	84
27	58	52	85
28	59	53	86
29	60	54	87
30	61	55	88
31	62	56	89
32	64	57	90
33	65	58	91
34	66	59	92
35	67	60	93
36	68		

## Acknowledgments

We acknowledge the data collection support from Sandra Applebaum and Roz Pierson at Harris Interactive. We thank Patrick Orr for his efforts in manuscript formatting, and researchers Omar Al-Shahery, Samuel Berkowitz, Dolly Dahdal, and Chaoling Feng at RAND for their efforts to review the Arabic and Chinese versions of the scale. Finally, we thank the many individuals who have participated in the cognitive interviews for their insight, which strengthened our research.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This work was funded by an unrestricted grant from the LIVESTRONG Foundation.

## References

- Ader, D. N. (2007). Developing the patient-reported outcomes measurement information system (PROMIS). *Medical Care*, 45(5), S1-S2.
- Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., . . . Highberger, L. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology*, 72, 105-118.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57, 289-300.
- Brewster, M. E., & Moradi, B. (2010). Perceived experiences of anti-bisexual prejudice: Instrument development and evaluation. *Journal of Counseling Psychology*, 57, 451-468. doi:10.1037/A0021116
- Brohan, E., Slade, M., Clement, S., & Thornicroft, G. (2010). Experiences of mental illness stigma, prejudice and discrimination: A review of measures. *BMC Health Services Research*, 10, Article 80. doi:10.1186/1472-6963-10-80
- Browne, M. W., Cudeck, R., Bollen, K. A., & Long, J. S. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, 154, 136.
- Cai, L., du Toit, S. H. C., & Thissen, D. (2011). *IRTPro version 2: Flexible, multidimensional, multiple categorical IRT modeling*. Chicago, IL: Scientific Software International.
- Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., . . . Rose, M. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. *Medical Care*, 45(5 Suppl. 1), S3-S11.
- DeWalt, D. A., Rothrock, N., Yount, S., & Stone, A. A. (2007). Evaluation of item candidates: The PROMIS qualitative item review. *Medical Care*, 45(5 Suppl. 1), S12-S21.
- Edelen, M. O., Stucky, B., & Chandra, A. (2013). Quantifying "problematic" DIF within an IRT framework: Application to a Cancer Stigma Index. *Quality of Life Research*. Advance online publication.
- Evans-Lacko, S., Little, K., Meltzer, H., Rose, D., Rhydderch, D., Henderson, C., & Thornicroft, G. (2010). Development and psychometric properties of the mental health knowledge schedule. *Canadian Journal of Psychiatry/Revue Canadienne De Psychiatrie*, 55, 440-448.
- Fries, J. F., Bruce, B., & Cella, D. (2005). The promise of PROMIS: Using item response theory to improve assessment of patient-reported outcomes. *Clinical and Experimental Rheumatology*, 23(5), S53-S57.
- Green, S., Davis, C., Karshmer, E., Marsh, P., & Straight, B. (2005). Living stigma: The impact of labeling, stereotyping, separation, status loss, and discrimination in the lives of individuals with disabilities and their families. *Sociological Inquiry*, 75, 197-215.
- Greene, K., & Banerjee, S. C. (2006). Disease-related stigma. *Journal of Homosexuality*, 50(4), 185-209.
- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1-55.
- Joachim, G., & Acorn, S. (2000). Living with chronic illness: The interface of stigma and normalization. *The Canadian Journal of Nursing Research/Revue canadienne de recherche en sciences infirmieres*, 32(3), 37-48.
- LIVESTRONG Foundation. (2007). *Cancer stigma and silence around the world*. Austin, TX: LIVESTRONG Foundation.
- Muthén, L. K., & Muthén, B. O. (2007). *1998-2007. Mplus user's guide*. Los Angeles, CA: Author.
- Pettit, M. L. (2008). Disease and stigma: A review of literature. *Health Educator*, 40(2), 70-76.
- Reeve, B. B., Hays, R. D., Bjorner, J. B., Cook, K. F., Crane, P. K., Teresi, J. A., . . . Cella, D. (2007). Psychometric evaluation and calibration of health-related quality of life item banks: Plans for the patient-reported outcomes measurement information system (PROMIS). *Medical Care*, 45(5 Suppl. 1), S22-S31. doi:10.1097/01.mlr.0000250483.85507.04
- Sartorius, N. (2007). Stigma and mental health. *The Lancet*, 370, 810-811.
- Thissen, D., Pommerich, M., Billeaud, K., & Williams, V. S. L. (1995). Item response theory for scores on tests including polytomous items with ordered responses. *Applied Psychological Measurement*, 19, 39-49.
- Uys, L. R., Holzemer, W. L., Chirwa, M. L., Dlamini, P. S., Greeff, M., Kohi, T. W., . . . Naidoo, J. R. (2009). The development and validation of the HIV/AIDS stigma instrument-nurse (HASI-N). *AIDS Care*, 21, 150-159.

## Author Biographies

**Maria Orlando Edelen**, PhD, is a senior Behavioral Social Scientist scientist and psychometrician at the RAND Corporation. Her research focuses primarily on the use of item response theory for development, refinement and evaluation of behavioral health measures.

**Anita Chandra**, PhD is a senior policy researcher and director of the Behavioral and Policy Sciences Department at the RAND Corporation. Her background is in public health, child and adolescent development, and community-based participatory research and evaluation. She currently leads or co-leads studies on community well-being; deployment and military families; community



resilience and long-term disaster recovery; and child health and development.

**Brian Stucky** is an Associate Behavioral Social Scientist at RAND Corporation, where he specializes in psychometrics and health research. He is a trained psychometrician whose research interests focus on developing and applying new test theory models (including factor analytic and item response theory-based) that serve to better understand patient response data in settings where multiple dimensions are present.

**Rebekkah M. Schear**, MIA has spent her career working at the intersection of global health and international development. Over the last five years at the LIVESTRONG Foundation, Rebekkah led the Foundation's implementation of six international programs throughout Latin America, Africa and Asia focusing on reducing cancer stigma and empowering cancer survivors to become advocates. Currently, she is working to create a new, innovative model of patient-centered cancer care with the goal of improving delivery of patient-centered cancer care globally.

**Claire Neal**, MPH is the Executive Director of the Triangle Global Health Consortium, a non-profit member organization representing institutions and individuals from the pharmaceutical and biotechnology industry, the international health development community, and academia – all dedicated to improving the health of the world's communities. Claire has over 13 years of experience in the development and delivery of innovative health programs. For the last decade, Claire has worked at the LIVESTRONG Foundation; most recently as the Vice President of Global Strategy where she provided the leadership and vision for LIVESTRONG's global work.

**Ruth Rechis**, PhD, is the Vice President of Programs & Strategy at the LIVESTRONG Foundation. As a member of the Executive Leadership Team, she oversees the day to day operations of the Foundation's programmatic work and shepherds the strategic planning process for the Foundation. As an adolescent cancer survivor, Dr. Rechis has a personal connection to the mission of the Foundation.