Psychological Test and Assessment Modeling, Volume 58, 2016 (1), 99-117

The Measuring Your Health study: Leveraging community-based cancer registry recruitment to establish a large, diverse cohort of cancer survivors for analyses of measurement equivalence and validity of the Patient Reported Outcomes Measurement Information System® (PROMIS®) short form items

Roxanne E. Jensen[1,2,10], *Carol M. Moinpour*[3], *Theresa H.M. Keegan*[4], *Rosemary D. Cress*[5,6], *Xiao-Cheng Wu*[7], *Lisa E. Paddock*[8,9], *Antoinette M. Stroup*[9] & *Arnold L. Potosky*[2,10]

Abstract

The Measuring Your Health (MY-Health) study was designed to fill evidence gaps by validating eight Patient Reported Outcomes Measurement Information System® (PROMIS®) domains (Anxiety, Depression, Fatigue, Pain Interference, Physical Function, Sleep Disturbance, Applied Cognitive Function, and Ability to Participate in Social Roles and Activities) across multiple race-ethnic

[1] *Correspondence concerning this article should be addressed to:* Roxanne E. Jensen, Ph.D., Cancer Prevention and Control Program, Lombardi Comprehensive Cancer Center, Georgetown University, 3300 Whitehaven Street NW, Suite 4100, Washington, DC 20007, USA; email: rj222@georgetown.edu

[2] Cancer Prevention and Control Program, Lombardi Comprehensive Cancer Center, Washington, DC

[3] Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA

[4] Department of Internal Medicine, Division of Hematology and Oncology, University of California Davis School of Medicine, Sacramento, CA

[5] Public Health Institute, Cancer Registry of Greater California, Sacramento, CA

[6] Department of Public Health Sciences, School of Medicine, University of California, Davis, CA

[7] LSU Health Sciences Center School of Public Health, Louisiana Tumor Registry, New Orleans, LA

[8] Cancer Epidemiology Services, New Jersey Department of Health, Trenton, NJ

[9] Rutgers School of Public Health and Cancer Institute of New Jersey, New Brunswick, NJ

[10] Department of Oncology, Georgetown University, Washington DC, USA

and age groups in a diverse cohort of cancer patients. This paper provides detailed information on MY-Health study design, implementation, and participant cohort; it identifies key challenges and benefits of recruiting a diverse community-based cancer cohort. Between 2010 and 2012, we identified eligible patients for the MY-Health study in partnership with four Surveillance, Epidemiology, and End Results (SEER) program cancer registries located in California, Louisiana, and New Jersey. The overall response rate for the MY-Health cohort (n = 5,506) was 34 %, with a median response time of 9.5 months after initial cancer diagnosis. The cohort represented meaningful diversity of age (22 % under 49 years of age) and race/ethnicity (41 % non-Hispanic White) across seven cancers. Challenges included lower response rates by race/ethnic minorities, young, and advanced-stage cancer patients, use of non-final registry information for eligibility identification, and lower use of translated surveys than expected. The MY-Health cohort represents one of the largest efforts to measure the full range of patient-reported symptoms experienced after initial cancer treatment. It provides sufficient diversity in terms of sociodemographics, symptoms, and function to provide a meaningful validation of eight PROMIS measures.

Key words: PROMIS, MY-Health, measurement equivalence, validity, cancer

Introduction

Studies evaluating patient-reported outcomes (PROs) in patients with many chronic health conditions, including cancer, have identified significant differences in PROs by age and across race/ethnic groups (Angel & Thoits, 1987; Osmond, Vranizan, Schillinger, Stewart, & Bindman, 1996; Raczynski et al., 1994; Shetterly, Baxter, Mason, & Hamman, 1996; Stewart & Napoles-Springer, 2003). However, the extent to which these variations reflect true differences or measurement bias remain unclear (Fullerton, Wallace, & Concha-Garcia, 1993; Skinner, Teresi, Holmes, Stahl, & Stewart, 2001; Teresi & Holmes, 1994). Differential Item Functioning (DIF), a type of measurement bias, occurs when individuals in different groups, such as race or age, respond differently to an item within a unidimensional measure, while reporting the same overall score or trait. DIF can affect the overall interpretations of PRO constructs by age, race/ethnicity and gender (Edwards, Doleys, Fillingim, & Lowery, 2001; Ibrahim, Burant, Mercer, Siminoff, & Kwoh, 2003; Sheffield, Biles, Orom, Maixner, & Sheps, 2000; Weiss, Emanuel, Fairclough, & Emanuel, 2001). Studies of commonly-administered generic and disease-specific PRO measures suggest that DIF is likely responsible for some of the observed group differences in both cancer and general populations (Crane, Gibbons, Narasimhalu, Lai, & Cella, 2007; Fleishman & Lawrence, 2003; Hahn et al., 2005; Teresi, Ramirez, Lai, & Silver, 2008). Therefore, while PROs have gained increasing recognition as legitimate endpoints in the evaluation of medical interventions' effects on function and well-being (Clauser, Ganz, Lipscomb, & Reeve, 2007; Ganz & Gotay, 2007), it is important to evaluate PRO measures for DIF to ensure their validity when administered within and across diverse populations. Given the expanding cultural diversification of the US population (Humes, Jones, & Ramirez, 2011), establishing the validity of PRO measures to accurately examine constructs across broad heterogeneous populations takes on increasing relevance.

In 2004, the National Institutes of Health launched the Patient Reported Outcomes Measurement Information System® (PROMIS®) "Roadmap Initiative" to use modern psychometric techniques to improve the measurement of symptoms and health outcomes by building and evaluating item banks from common, accessible tools (Cella et al., 2007). This initiative created PRO measures covering a wide range of symptoms and function, establishing a standardized scoring framework that could be used across illnesses, chronic health conditions, and the general population. Initial validity and reliability efforts for PROMIS® measures rarely included enough racially and ethnically diverse patients to establish the validity of these measures for the U.S. population. The Measuring Your Health (MY-Health) study was designed to fill this evidence gap and evaluate eight PROMIS domains across multiple race-ethnic and age groups in a diverse cohort of cancer patients. It has accomplished this by partnering with four Surveillance, Epidemiology, and End Results (SEER) program cancer registries to draw a population-based sample of recently diagnosed cancer patients, and oversampling race/ethnic minorities and younger patients. This collaboration has allowed the MY-Health study cohort to provide an extensive and generalizable cross-cultural validation of the PROMIS measures in a large community-based sample. The goal of this paper is to provide detailed information on the MY-Health study design, implementation, and participant cohort used in the PROMIS validation papers presented in this special issue. It also discusses the challenges and benefits of recruiting and enrolling a diverse community-based cancer cohort.

Methods

Identification and Recruitment. Between 2010 and 2012, we identified eligible patients for the MY-Health study in partnership with four SEER cancer registries located in California (two), Louisiana, and New Jersey. We selected SEER registry sites for two reasons: to represent the diversity of the U.S. population with respect to age, sex, race-ethnicity, and socioeconomic status, and to recruit study participants from regions of the country that represent a wide spectrum of cultures, access to care, and emigration countries of origin.

The SEER registry sites identified patients based on eligibility criteria, and then mailed self-administered surveys along with a cover letter containing IRB-required language regarding the purpose of the study and the voluntary nature of their participation. In Louisiana, patient physicians were notified first, allowing for opt-out due to medical reasons. Spanish and Mandarin (traditional and simplified characters available) language surveys, with cover letters in the same language were included along with the English survey in the initial mailings to eligible participants based on race/ethnicity, or made available on request. For all SEER registries race-ethnicity identification was collected from the medical record supplemented with linkage to the NAACCR Hispanic and Asian/Pacific Islander Identification Algorithm (NHAPIIA). This algorithm uses gender, surname and birthplace to better identify Hispanics and Asian/Pacific Islanders (NAACCR Latino Research Work Group, 2005; NAACCR Race and Ethnicity Work Group, 2011).

All non-responders received a second mailing of the patient survey at three weeks after the first mailing. Following another three weeks, phone follow-up (English, Spanish, Mandarin language options available) was initiated for all non-responders to answer questions about the study, encourage participation or offer completion of the survey over the phone. Eligible participants who were unable to be reached after five phone attempts at different time slots (day and evening time) and days (weekday and weekend) were considered passive refusers. Participants received a $30 gift card or check after completing the baseline survey. Participants also completed a six month follow-up survey, and we conducted a detailed medical record abstraction from a random sub-sample (participant follow-up is not discussed further in this paper).

Eligible Population. Eligible participants were 21-84 years old at the time of initial diagnosis of their first primary cancer. We restricted survey eligibility to persons diagnosed with one of seven cancers (prostate, colorectal, non-small cell lung, Non-Hodgkin lymphoma, female breast, uterine or cervical) between six to thirteen months of diagnosis, and able to read English, Spanish, or Mandarin. Sampling was stratified by four race-ethnicity groups (Non-Hispanic White [NHW], Hispanic, Non-Hispanic Black [Black], Non-Hispanic Asian/ Pacific Islanders [Asian]) and three age groups at diagnosis (21-49, 50-64, 65-84 years). A goal was to have approximately 1000 in each of the ethnic groups to permit latent variable modeling of the PROMIS items.

We chose cancer types to facilitate validation and reflect a wide range of symptoms and functions, and ensure a sufficient number of younger (21-49) participants. The time period (six to thirteen months) was selected to allow enough time for each SEER registry to identify and verify the eligibility of new cancer cases. The seven month recruitment period allowed for multiple survey mailings and phone follow-up. This study was approved by Institutional Review Boards at Georgetown University, the State of California, and all participating research sites.

Language Translations. The full MY-Health survey was translated into Mandarin (traditional and simplified) and Spanish. When available, the official PROMIS measure translations were used; otherwise, a new translation was performed. All MY-Health survey translations (of PROMIS, other PROs, and all survey text) were performed using PROMIS translation methodology and procedures (Correia, 2013), which have been adapted from FACIT translation methodology (Eremenco, Cella, & Arnold, 2005). It is an iterative process with the goal of producing one global language version suitable for all countries where the language is spoken. The procedures include two forward translations by native Mandarin-speaking professionals, one English back-translation by an English-speaking translator, review by three bilingual experts, and cognitive testing with at least five native speakers. The MY-Health translations were conducted using cognitive testing with 30 people (10 native Spanish-speaking Hispanic cancer survivors, 10 native Mandarin-speaking Chinese participants [4 non-cancer, 6 cancer survivors], and 10 native English-speaking Black and NHW cancer survivors). The translation process was coordinated with the PROMIS Statistical Center at Northwestern University. This methodology has been recently applied and validated in a Dutch-Flemish translation of 17 PROMIS domains (Terwee et al., 2014).

Survey Measures. We chose eight PROMIS domains for inclusion in the MY-Health study based on their prevalence and impact in cancer patient populations: Emotional Distress – Anxiety (11-item, Cronbach's alpha [α] = 0.96); Emotional Distress – Depression (10-item, α =0.96); Fatigue (14-item, α =0.95); Pain Interference (11 items, α = 0.98); Physical Function (16-item, α =0.96), Sleep Disturbance (10-item, α =0.94), Applied Cognition – General Concerns (8-item, α = 0.97); and Ability to Participate in Social Roles and Activities v.2 (10-item, α = 0.97). All PROMIS measures, except Ability to Participate in Social Roles and Activities and Applied Cognition – General Concerns, were normalized to the general US population (Cella et al., 2007). We used custom short forms to measure these domains because of our emphasis on including as many items as possible for validation analysis. Selection of items for each domain was based on their inclusion on short forms (as of 2010) or their high frequency of selection when administered online using the PROMIS computerized adaptive testing (CAT) through assessment center. The latter assessment of high frequency CAT items was based on a prior sample of cancer patients scoring at least one-half standard deviation above (i.e., higher symptoms) than the US general population mean (See Table 1).

We also included the following PRO measures: A legacy measure of Physical Well-being, the FACT-G Physical Well-Being (PWB) subscale (α = 0.90; Cella et al., 1993); Spirituality, comprised of two sub-domains (faith and peace) measured by the FACIT-SP-12 v4 (α = 0.81; Peterman, Fitchett, Brady, Hernandez, & Cella, 2002); Financial

Table 1:
PROMIS short form coverage in the Measuring Your Health (MY-Health) study by domain

Domain	**PROMIS Short Form Coverage**							**Number of MY-Health Survey items (Custom)**	**Total PROMIS items available in item bank**
	4a	**6a**	**6b**	**7a**	**8a**	**8b**	**10a**		
Pain – Interference			x					10	40
Fatigue	x	x		x				14	95
Emotional Distress – Depression	x	x				x		10	28
Emotional Distress – Anxiety	x	x		x	x			11	22
Sleep Disturbance	x	x	x			x		10	27
Ability to Participate in Social Roles and Activities (v2)								10	35
Physical Function	x	x	x				x	16	121
Applied Cognition – General Concerns (v2)	x	x						8	71

Well-being subscale from the PSQ-III (4-item, adapted, α = 0.82; Ware, Davies-Avery, & Stewart, 1978); and U.S. acculturation (adapted to reflect English vs. any other language, α = 0.93; Marin, Sabogal, Marin, Otero-Sabogal, & Perez-Stable, 1987). We also used a single-item Patient Self-Report ECOG Performance Status Scale utilized in cancer clinical trials to assess disease impact on daily living (Oken et al., 1982). Clinically moderate and severe symptom thresholds reported for pain, anxiety, depression, and fatigue are defined elsewhere (Cella et al., 2014).

Finally, the MY-Health survey collected self-reported race, ethnicity, education level, current employment status, annual household income, marital status, health insurance coverage, acculturation (born in US, years in US, and the Marin U.S. acculturation scale, described above), receipt of cancer treatments, comorbidities, and selected health behaviors (e.g., weekly exercise, smoking status). After data collection, we grouped persons by race-ethnicity according to the U.S. Census (2010) classification: White, Black or African American, American Indian or Alaska Native, Asian, and Native Hawaiian or Other Pacific Islander, as well as, Hispanic or Latino and Not Hispanic or Latino (Humes et al., 2011). Other self-identified races not captured in this classification, or the selection of multiple races was also included.

Registry Data. We obtained SEER registry data to enrich our study dataset with variables not feasibly obtained via patient survey. We collected age at diagnosis, sex, date of cancer diagnosis, cancer site, cancer stage, and first course of treatment from the cancer registries. Cancer site specific variables (e.g., results from a HER2: Immunohistochemistry (IHC) Test for breast cancer) were also obtained from the registries. All information from the registries was merged with survey data for each MY-Health study participant at the Georgetown coordinating center.

Results

The baseline survey was completed by 5,506 cancer patients (See Figure 1). Table 2 shows the demographic characteristics and Table 3 shows the clinical characteristics of the overall study cohort. The ethnic/racial group designations used in the analyses reflect "gold-standard" self-reported race/ethnicity as shown in Table 2. Fifty-Nine percent of participants were under 65 years of age, and 41 % were White. Asian and Hispanic patients were the least likely to report being born in the U.S. (16 % and 41 %, respectively; see Table 2), representing diversity by country of origin (Table 4). Over half of the participants (51 %) reported an income of less than $60,000 a year, and 37 % reported a high school degree or less. Breast, prostate, and colorectal were the most common cancers (30 %, 21 % and 17 %, respectively), and 67 % of all participants were diagnosed with either stage I or II cancer (Table 3). According to self-reported performance status, close to half the cohort reported "no symptoms" (45 %), while 16 % reported "being on bed rest" for at least some part of the day. Multiple (two or more) comorbidities were identified by 40 % of the cohort. Moderate or higher fatigue was the most common patient-reported PROMIS symptom (41 %), followed by pain interference (31 %).

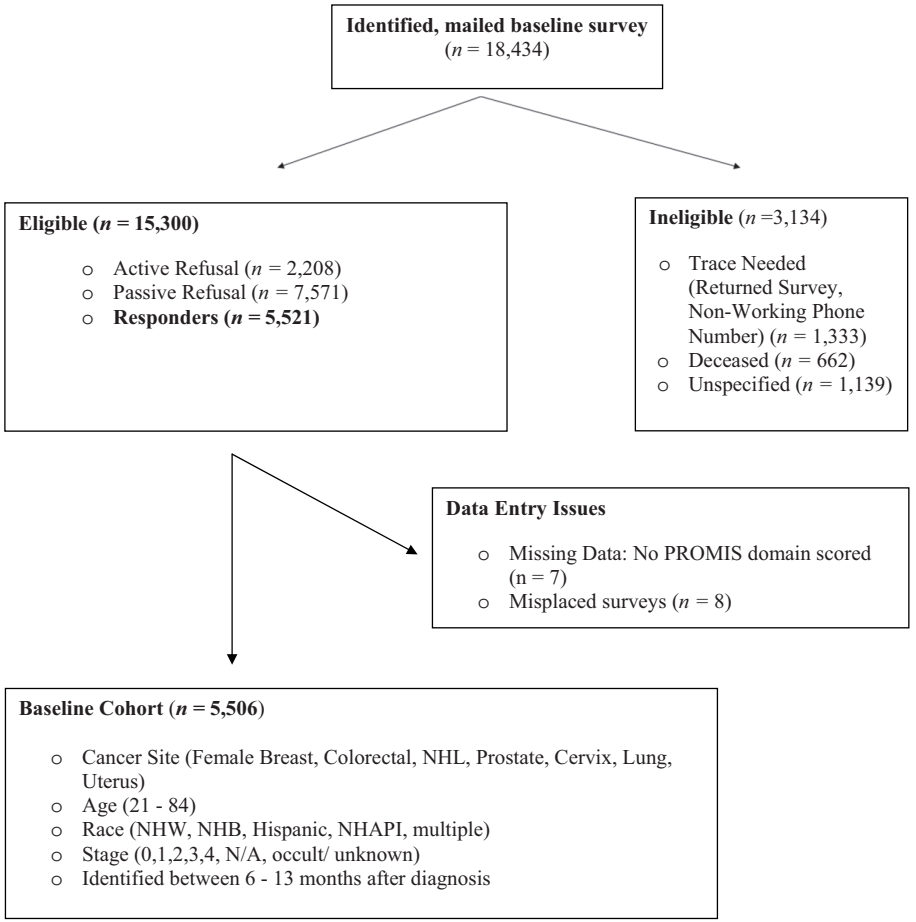

Figure 1:
Measuring Your Health (MY-Health) Survey Recruitment Flow Chart

Table 2:
Demographic characteristics of the Measuring-Your Health study cohort (n = 5506)

Demographic characteristics		
	Overall	
	n	**%***
Age at Diagnosis (years)		
21-49	1,203	22
50-64	2,037	37
65-84	2,266	41
Sex		
Male	2,222	40
Female	3,284	60
SEER Region		
Greater California	1,864	34
Greater Bay Area (San Francisco)	1,254	23
Louisiana	1,086	20
New Jersey	1,302	24
Married	3,200	58
Education Level		
< High School Degree	981	18
High School Degree	1,061	19
Some College	1,766	32
College Degree	981	18
Graduate Degree	641	12
Missing/Unknown	76	1
Income Level		
< $10,000	584	11
$10,000 to $59,999	2,186	40
$60,000 to $99,999	908	16
$100,000 to $199,999	674	12
> $200,000	189	3
Missing/Unknown	965	18
Employment Status		
Working	2,377	43
Retired	2,114	38
Unemployed/Disabled	933	17
Missing	82	1

Health Insurance Coverage		
Private	2,274	41
Government	1,631	30
Private & Government	1,317	24
No Insurance	116	2
Missing/Unknown	168	3
Born in U.S.	3,854	70
Survey Language		
English	5,011	91
Spanish	352	6
Chinese	143	3
Survey Administration Mode		
Paper	5,408	98
Phone	98	2
Race/Ethnicity		
White	2,261	41
Black	1,121	20
Hispanic	1,064	19
Asian	887	16
*Other***	28	1
Multiple	145	3

*Due to missing values numbers may not equal 100 %
** Other Race: Alaska Native/American Indian, Asian Hawaiian/Pacific Islander, Self-Identified "Other"

Table 3:
Clinical characteristics of the Measuring-Your Health study cohort

Clinical Characteristics		
	Overall	
	n	**%***
Cancer Type		
Breast	1,662	30
Cervix	149	3
Colorectal	937	17
Lung	722	13
NHL	464	8
Prostate	1,177	21
Uterus	395	7
Stage at Diagnosis		
I	1,983	36
II	1,731	31
III	935	17
IV	635	12
*Missing/Unknown***	222	4
Comorbidities (Number)		
0	1,920	35
1	1,394	25
2+	2,192	40
Initial Treatment Type***		
Surgery	3,748	68
Chemotherapy	2,642	48
Hormonal Therapy	1,208	22
Radiation	2,264	41
Moderate to Severe Symptoms***		
Anxiety	521	9
Depression	446	8
Pain Interference	1,683	31
Fatigue	2,257	41
Performance Status		
No Symptoms	2,465	45
Some Symptoms	2,054	37
<50 % Bed Rest	681	12
>50 % Bed Rest	220	4
Unable to get out of bed	23	0
Missing	63	1

*due to missing values numbers may not equal 100 %
** Includes: In situ, Occult, N/A staging reported by registry
*** Categorical options are not mutually exclusive

Table 4:
Country of Origin and English Language

Country of Origin	**Total**		**English Surveys**	
	n	**%**	***n***	**%**
ASIAN				
Chinese	310	35	183	59
Filipino	256	29	256	100
Japanese	80	9	80	100
Asian Indian	75	8	72	96
Vietnamese	61	7	59	97
Korean	34	4	34	100
Other Asian (Write-in)	45	5	45	100
Multiple Selected	19	2	15	79
Unknown	7	1	7	100
TOTAL	**887**	**100**		
HISPANIC				
Mexican, Mexican American, Chicano	578	54	389	67
Puerto Rican	94	9	77	82
Cuban	27	3	17	63
Dominican	28	3	7	25
Other Hispanic (Write-in)	320	30	208	65
Columbian	23	2	14	60
Salvadorian	24	2	5	21
Peruvian	19	2	5	26
Guatemalan	16	2	3	19
Ecuadorian	15	1	4	27
Nicaraguan	12	1	6	50
Other	211	20	171	81
Unknown	17	2	16	94.1
TOTAL	**1,064**	**100**		

Table 5:
Response rates and eligibility status of the MY-Health cohort by Overall Surveillance, Epidemiology and End Results (SEER) cancer registry variables (2010 - 2012)

	MY-Health Cohort								
	Identified		**Total Eligible**		**Final Cohort**			**Overall SEER (4 Study Registries, 2010, 2011, and 2012)**	
	n	**%***	**n**	**%****	***n***	**%****	**%***	***n***	**%***
Total	**18434**	**100**	**15300**	**83**	**5506**	**36**	**100**	**209,419**	**100**
Race/Ethnicity‡									
White	7,640	41	6,559	86	2,606	40	47	138,543	66
Black	3,847	21	3,318	86	1,184	36	22	24,244	12
Hispanic	2,669	14	2,277	85	705	31	13	25,969	12
Asian	2,818	15	2,321	82	738	32	13	14,183	7
Other‡‡	971	5	825	85	273	33	5	6,480	3
Missing	489	3	0	0	0	0	0	0	0
Age at Diagnosis									
21 - 49	3,512	19	3,085	88	1,203	39	22	27,081	13
50 - 64	6,555	36	5,632	86	2,037	36	37	83,832	40
65 - 84	7,871	43	6,583	84	2,266	34	41	98,506	47
>84	2	0	0	0	0	0	0	0	0
Missing	494	3	0	0	0	0	0	0	0
Stage at Diagnosis									
I	6,039	33	5,438	90	1,983	36	36	71,160	34
II	5,022	27	4,505	90	1,731	38	31	60,337	29
III	2,989	16	2,514	84	935	37	17	28,048	13
IV	2,921	16	2,069	71	635	31	12	37,598	18
Other‡‡‡	1,463	8	774	53	222	29	4	12,276	6
Cancer Site									
Breast	4,476	24	4,074	91	1,662	41	30	57,806	28
Cervix	551	3	447	81	149	33	3	3,825	2
Colorectal	3,233	18	2,732	85	937	34	17	14,480	7
Lung	2,950	16	2,142	73	722	34	13	41,285	20
NHL	1,603	9	1,331	83	464	35	8	15,570	7
Prostate	3,736	20	3,392	91	1,177	35	21	63,273	30
Uterus	1,342	7	1,182	88	395	33	7	13,180	6
Missing	543	3	0	0	0	0	0	0	0
Sex									
Male	7,787	42	6,538	84	2,223	34	40	101,664	49
Female	10,158	55	8,762	86	3,283	37	60	107,755	51
Missing	489	3	0	0	0	0	0	0	0

* Column %
** Row %
‡ Registry Derived Race/Ethnicity using gender, surname, and birthplace
‡‡ American Indian, Alaska Native, Hawaiian, Pacific Islander, Other, Multiple
‡‡‡ In situ, N/A, Occult, Unknown, Missing

We obtained an overall response rate for the MY-Health cohort of 36 %, with the median response time of 9.5 months (range 6 - 13) after the initial cancer diagnosis. Of all patients identified by the SEER registries by our preliminary eligibility criteria, 83 % remained eligible after final SEER clinical data verification. Among eligible participants, 49 % were passive refusals (working number and address, no response after at least four phone attempts), and 14 % were active decliners (reached by study team member over phone). Table 5 presents study response rates by registry-derived patient characteristics. (As previously noted, the sample sizes for reported race/ethnicity in Table 5 are lower than those used in the analyses because the self-reported designation was used.) We found that eligible patient response rates were significantly (all $p < 0.001$) higher among patients who were White (40 % versus 31-36 % for race-ethnic minorities), younger (21-49 years, 39 % versus 65 – 84, 41 %) or diagnosed with non-metastatic cancer (36 - 38 %, versus 29 % for participants diagnosed with stage IV cancer; Table 5). Among eligible participants, passive refusal was highest among Blacks and Hispanics (53 % and 58 %, respectively), and active refusal was highest among Asians (19%). Among all identified patients, lack of current contact information was highest for Asians (10 %), patients 21-49 years of age (9 %), and patients with advanced stage at diagnosis (11 %).

Due to the study sampling, the MY-Health participants were not representative of the overall SEER population, and included higher proportions of younger and non-White participants (Table 5). In contrast, differences by cancer type and stage at diagnosis between MY-Health participants and the overall SEER population are minimal.

Discussion

Overall, this study demonstrates a large-scale recruitment of a diverse cancer patient cohort through multiple SEER registries. The MY-Health cohort represents one of the largest efforts to measure the full range of patient symptoms experienced after most initial cancer treatment has been completed. It provides sufficient diversity in terms of sociodemographics, symptoms, and function to provide a meaningful validation of eight PROMIS measures covering the full PROMIS adult self-reported health domain framework (physical, mental, and social domains; PROMIS Network, 2015a). It also provides sufficient sample sizes to test for DIF in PROMIS measures with respect to age, race/ethnicity, and other important a priori patient sub-groups. Furthermore, this is a community-based sample, ensuring information on symptoms and function represents cancer patients who may have limited access to medical care.

The overall response rate for this cohort is low; however, the response rate among patients who were in contact with study staff was higher (71 %, excluding both ineligible and eligible passive refusal). These rates are consistent with similar cancer registry-based surveys (Arora, Reeve, Hays, Clauser, & Oakley-Girvan, 2011; Catalano et al., 2013; Harlan et al., 2011). Another consideration when interpreting these response rates was that we targeted non-White ethnic/minority participants, and patients with metastatic disease. Our findings are supported by research suggesting that race-ethnic minorities are less likely to participate in health surveys (Yancey, Ortega, & Kumanyika, 2006), re-

flecting a known recruitment issue. In addition, studies have also documented the difficulty recruiting participants with a recently diagnosed terminal prognosis (Addington-Hall, 2002).

This study also identified that younger patients, advanced-stage patients, and specific ethnic/minority groups (Asian, Hispanic) were less likely to have current contact information (phone or address), eliminating any possible outreach. Our ability to contact a higher proportion of older, non-Hispanic White Americans is supported by data showing that this group may be less transitory and more likely to be home owners (Boehm & Schlottmann, 2004; U.S. Census Bureau, 2012). Patients diagnosed with advanced-stage cancer (such as lung) may, in contrast, be transitioning to other living situations, such as moving in with family, assisted living, or hospice (Kutner & Kilbourn, 2009). While SEER registries devote considerable effort to tracking and updating patient contact information, this recruitment barrier illustrates the wide range of underlying difficulties in study enrollment.

While these participation differences could represent a meaningful study bias, the benefits of a community-based patient sample and degree of clinical information available from SEER-linked patient identification and recruitment are noteworthy. This study design provides a degree of population diversity not typically seen in measurement validation. A better understanding of both participant identification and recruitment issues linked with participation barriers could help future studies better target this patient population.

One of the most challenging recruitment issues was identifying, contacting, and having completed surveys returned within a small time window close to the initial diagnosis date (within six to thirteen months of identification). Because final SEER data were not available, 6 % of eligible patients, based on initial sampling criteria (one of seven cancers, first cancer diagnosis, within our pre-specified age and race/ethnicity sampling stratification groups), were deemed ineligible after completing the survey. Common reasons for these exclusions included updated SEER registry information that resulted in ineligibility, such as being previously diagnosed with a different cancer, or having a revised date of diagnosis falling outside the eligibility window. While this was a challenge to our data collection efforts, the SEER registries were constantly working to update and finalize their patient information, allowing for a high degree of confidence in the cancer clinical variables captured for this study cohort.

An unexpected issue was the limited use of our Spanish- and Chinese-translated surveys by participants. One reason for the limited use of the Chinese translation could have been partially due to the high representation of Asian participants from English-speaking countries (e.g., Philippines and India), suggesting that the limited use of the Chinese translation does not alone reflect a data quality issue. However, it does suggest additional translations are needed to adequately survey non-English speaking Asian immigrants. A separate reason for the lower use of translated surveys may have been our initial approach. At first we provided surveys in Spanish or Chinese upon request, rather than including translated surveys and cover letters in our initial mailings. After a 6-month recruitment review, we changed our procedure to send translated surveys in the initial

mailing based on SEER race/ethnic identification. This change increased the use of translated surveys, but also increased the mailing costs and volume of study materials sent to participants. As health surveys migrate to electronic/web-based administration, a strong benefit will be the improved ability to offer real-time translations to accommodate a diverse patient sample. PROMIS measures offer a wide number of translations, all following the same translation methodology (PROMIS Network, 2015b).

While these recruitment issues were meaningful operational challenges for our study team, the resulting MY-Health cohort provides extraordinary demographic variation in socioeconomic status, alongside important verified clinical information and a wide range of both symptom severity and functional disability. While our cohort by design is not representative of the SEER population, a cohort of this size is powered to identify meaningful DIF, ultimately supporting the validity of PROMIS measures across age and race/ethnic groups.

Funding

U01AR057971 (PIs: Potosky, Moinpour), NCI P30CA051008 (Jensen, Potosky), KL2TR000102 from the National Center for Research Resources (NCRR), and the National Center for Advancing Translational Sciences (NCATS), and National Institutes of Health (NIH), through the Clinical and Translational Science Awards Program (CTSA; Jensen).

The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement U58DP003862-01 awarded to the California Department of Public Health.

The ideas and opinions expressed herein are those of the author(s). It does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. An endorsement by the State of California, Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors is not intended nor should be inferred.

Acknowledgements

Georgetown MY-Health coordination center staff (Tania Lobo, Grace Zhou, Charlene Kuo, Lindsay Wright, Caroline Moore, Marin Rieger, Aaron Roberts, Deena Loeffler) and SEER MY-Health site collaborators (Laura Allen, Lauren S. Maniscalco, Lisa Moy, Natalia Herman, Wendy Ringer, Dr. Kevin Henry, and Dr. John Graff).

References

Addington-Hall, J. (2002). Research sensitivities to palliative care patients. *European Journal of Cancer Care, 11*(3), 220-224. doi: 10.1046/j.1365-2354.2002.00343.x

Angel, R., & Thoits, P. (1987). The impact of culture on the cognitive structure of illness. *Cultural Medicine Psychiatry, 11*(4), 465-494. doi: 10.1007/BF00048494

Arora, N. K., Reeve, B. B., Hays, R. D., Clauser, S. B., & Oakley-Girvan, I. (2011). Assessment of quality of cancer-related follow-up care from the cancer survivor's perspective. *Journal of Clinical Oncology, 29*(10), 1280-1289. doi: 10.1200/JCO.2010.32.1554

Boehm, T. P., & Schlottmann, A. M. (2004). The dynamics of race, income, and homeownership. *Journal of Urban Economics, 55*(1), 113-130. doi:http://dx.doi.org/10.1016/j.jue.2003.08.001

Catalano, P. J., Ayanian, J. Z., Weeks, J. C., Kahn, K. L., Landrum, M. B., Zaslavsky, A. M.,…The Cancer Care Outcomes Research & Surveillance Consortium (2013). Representativeness of participants in the cancer care outcomes research and surveillance consortium relative to the surveillance, epidemiology, and end results program. *Medical Care, 51*(2), e9-15. doi: 10.1097/MLR.0b013e318222a711. doi: 10.1097/MLR.0b013e318222a711.

Cella, D. F., Choi, S., Garcia, S., Cook, K. F., Rosenbloom, S., Lai, J. S.,… Gershon, R. (2014). Setting standards for severity of common symptoms in oncology using the PROMIS item banks and expert judgment. *Quality of Life Research, 23*(10), 2651-2661. doi: 10.1007/s11136-014-0732-6

Cella, D. F., Tulsky, D. S., Gray, G., Sarafian, B., Linn, E., Bonomi, A.,… Brannon, J. (1993). The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *Journal of Clinical Oncology, 11*(3), 570-579.

Cella, D. F., 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. doi:10.1097/01.mlr.0000258615.42478.55.

Clauser, S. B., Ganz, P. A., Lipscomb, J., & Reeve, B. B. (2007). Patient-reported outcomes assessment in cancer trials: evaluating and enhancing the payoff to decision making. *Journal Clinical Oncology, 25*(32), 5049-5050. doi: 10.1200/JCO.2007.14.5888

Correia, H. (2013). PROMIS® Instrument Development and Validation Scientific Standards Version 2.0 *Translation & Cultural Adaptation: PROMIS Guideline Document-Appendix 5.* (pp. 22-26). Chicago, IL: PROMIS Cooperative Group.

Crane, P. K., Gibbons, L. E., Narasimhalu, K., Lai, J. S., & Cella, D. (2007). Rapid detection of differential item functioning in assessments of health-related quality of life: The Functional Assessment of Cancer Therapy. *Quality of Life Research, 16*(1), 101-114. doi: 10.1007

Edwards, R. R., Doleys, D. M., Fillingim, R. B., & Lowery, D. (2001). Ethnic differences in pain tolerance: clinical implications in a chronic pain population. *Psychosomatic Medicine, 63*(2), 316-323. doi: 10.1097/00006842-20010300-0018

Eremenco, S. L., Cella, D., & Arnold, B. J. (2005). A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Evaluation & the Health Professions, 28*(2), 212-232. doi: 28/2/212 [pii];10.1177/0163278705275342 [doi]

Fleishman, J. A., & Lawrence, W. F. (2003). Demographic variation in SF-12 scores: true differences or differential item functioning? *Medical Care, 41*(7 Suppl), III75-III86. doi: 10.1097/01.MLR.0000076052.42628.CF

Fullerton, J. T., Wallace, H. M., & Concha-Garcia, S. (1993). Development and translation of an English-Spanish dual-language instrument addressing access to prenatal care for the border-dwelling Hispanic women of San Diego County. *Journal of Nurse Midwifery, 38*(1), 45-50. DOI: 10.1097/01.MLR.0000076052.42628.CF

Ganz, P. A., & Gotay, C. C. (2007). Use of patient-reported outcomes in phase III cancer treatment trials: lessons learned and future directions. *Journal of Clinical Oncology, 25*(32), 5063-5069. doi: 10.1200/JCO.2007.11.0197

Hahn, E., Holzner, B., Kemmler, G., Sperner-Unterweger, B., Hudgens, S., & Cella, D. (2005). Cross-cultural evaluation of health status using Item Response Theory: FACT-B comparisons between Austrian and US patients with breast cancer. *Evaluation and the Health Professions, 28*, 233-259. doi: 10.1177/0163278705275343

Harlan, L. C., Lynch, C. F., Keegan, T. H., Hamilton, A. S., Wu, X. C., Kato, I.,… Group, A. H. S. C. (2011). Recruitment and follow-up of adolescent and young adult cancer survivors: the AYA HOPE Study. *Journal of Cancer Survivorship, 5*(3), 305-314. doi: 10.1007/s11764-011-0173-y

Humes, K. R., Jones, N. A., & Ramirez, R. R. (2011). *Overview of Race and Hispanic Origin: 2010*. U. S. Department of Commerce; Economic and Statistics Administration Retrieved from http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf.

Ibrahim, S. A., Burant, C. J., Mercer, M. B., Siminoff, L. A., & Kwoh, C. K. (2003). Older patients' perceptions of quality of chronic knee or hip pain: differences by ethnicity and relationship to clinical variables. *Journal of Gerontology A Biological Sciences Medical Sciences, 58*(5), M472-M477. DOI: 10.1093/gerona/58.5.M472

Kutner, J. S., & Kilbourn, K. M. (2009). Bereavement: Addressing Challenges Faced by Advanced Cancer Patients, Their Caregivers, and Their Physicians. *Primary Care: Clinics in Office Practice, 36*(4), 825-844. doi:http://dx.doi.org/10.1016/j.pop.2009.07.004. doi.org/10.1016/j.pop.2009.07.004.

Marin, G., Sabogal, F., Marin, B. V., Otero-Sabogal, R., & Perez-Stable, E. J. (1987). Development of a short acculturation scale for Hispanics. *Hispanic Journal of Behavioral Sciences, 9*(2), 183-205. doi: 110.1177/07399863870092005.

NAACCR Latino Research Work Group. (2005). NAACCR Guideline for Enhancing Hispanic-Latino Identification: Revised NAACCR Hispanic/Latino Identification Algorithm [NHIA v2] Springfield, IL North American Association of Central Cancer Registries

NAACCR Race and Ethnicity Work Group. (2011). NAACCR Asian Pacific Islander Identification Algorithm [NAPIIA v1.2.1]. Springfield, IL North American Association of Central Cancer Registries

Oken, M. M., Creech, R. H., Tormey, D. C., Horton, J., Davis, T. E., McFadden, E. T.,… Carbone, P. P. (1982). Toxicity and response criteria of the Eastern Cooperative Oncology Group. *American Journal of Clinical Oncology, 5*(6), 649-655. doi: 10.1097/00000421-198212000-00014

Osmond, D. H., Vranizan, K., Schillinger, D., Stewart, A. L., & Bindman, A. B. (1996). Measuring the need for medical care in an ethnically diverse population. *Health Service Research, 31*(5), 551-571.

Peterman, A. H., Fitchett, G., Brady, M. J., Hernandez, L., & Cella, D. (2002). Measuring spiritual well-being in people with cancer: the functional assessment of chronic illness therapy – Spiritual Well-being Scale (FACIT-Sp). *Annals of Behavioral Medicine, 24*(1), 49-58. doi: 10.1207/S15324796ABM2401_06

PROMIS Network. (2015a). Domain Frameworks: PROMIS Adult Self-Reported Health Retrieved June 19, 2015, from http://www.nihpromis.org/measures/domainframework1

PROMIS Network. (2015b). Translations Retrieved June 17 2015, from http://www.nihpromis.org/measures/translations

Raczynski, J. M., Taylor, H., Cutter, G., Hardin, M., Rappaport, N., & Oberman, A. (1994). Diagnoses, symptoms, and attribution of symptoms among black and white inpatients admitted for coronary heart disease. *American Journal of Public Health, 84*(6), 951-956. doi: 10.2105/AJPH.84.6.951

Sheffield, D., Biles, P. L., Orom, H., Maixner, W., & Sheps, D. S. (2000). Race and sex differences in cutaneous pain perception. *Psychosomatic Medicine, 62*(4), 517-523.

Shetterly, S. M., Baxter, J., Mason, L. D., & Hamman, R. F. (1996). Self-rated health among Hispanic vs non-Hispanic white adults: the San Luis Valley Health and Aging Study. *American Journal Public Health, 86*(12), 1798-1801. doi: 10.2105/AJPH.86.12.1798

Skinner, J. H., Teresi, J. A., Holmes, D., Stahl, S. M., & Stewart, A. (2001). Measurement in older ethnically diverse populations: Overview of the volume *Journal of Mental Health and Aging* (Vol. 7, pp. 5-8).

Stewart, A. L., & Napoles-Springer, A. M. (2003). Advancing health disparities research: can we afford to ignore measurement issues? *Medical Care, 41*(11), 1207-1220. doi: 10.1097/01.MLR.0000093420.27745.48

Teresi, J. A., & Holmes, D. (1994). Overview of methodological issues in gerontological and geriatric measurement. In L. MP & T. JA (Eds.), *Focus on assessment techniques* (pp. 1-22). New York, NY: Springer Publishing Company.

Teresi, J. A., Ramirez, M., Lai, J. S., & Silver, S. (2008). Occurrences and sources of Differential Item Functioning (DIF) in patient-reported outcome measures: description of DIF methods, and review of depression, quality of life and general health. *Psychology Science Quarterly, 50*(4), 538-612.

Terwee, C. B., Roorda, L. D., de Vet, H. C., Dekker, J., Westhovens, R., van leeuwen, J.,… Boers, M. (2014). Dutch-Flemish translation of 17 item banks from the Patient-Reported Outcomes Measurement Information System (PROMIS). *Quality of Life Research*. doi: 10.1007/s11136-013-0611-6

U.S. Census Bureau. (2012). Homeownership Rates by Age of Householder and Household Type: 1990 to 2010. *Statistical Abstract of the United States* Retrieved June 19 2015, from http://www.census.gov/compendia/statab/2012/tables/12s0992.pdf

Ware, J. E., Davies-Avery, A., & Stewart, A. L. (1978). The measurement and meaning of patient satisfaction. *Health & Medical Care Services Review, 1*(1).

Weiss, S. C., Emanuel, L. L., Fairclough, D. L., & Emanuel, E. J. (2001). *Lancet, 357*(9265), 1311-1315. doi:10.1016/S0140-6736(00)04515-3

Yancey, A. K., Ortega, A. N., & Kumanyika, S. K. (2006). Effective Recruitment and Retention of Minority Research Participants. *Annual Review of Public Health, 27*(1), 1-28. doi: doi:10.1146/annurev.publhealth.27.021405.102113.