



Published in final edited form as:

Amyloid. 2020 June ; 27(2): 111–118. doi:10.1080/13506129.2020.1713743.

The use of PROMIS patient-reported outcomes (PROs) to inform light chain (AL) amyloid disease severity at diagnosis

Anita D'Souza^{1,2}, Brooke E. Magnus³, Judith Myers², Angela Dispenzieri⁴, Kathryn E Flynn^{1,2}

¹Division of Hematology/Oncology, Department of Medicine, Medical College of Wisconsin, Milwaukee, WI

²Center for International Blood and Marrow Transplant Research

³Department of Psychology, Marquette University, Milwaukee, WI

⁴Division of Hematology, Department of Medicine, Mayo Clinic, Rochester, MN

Abstract

We sought to evaluate how PROMIS patient-reported outcome (PRO) measures correlated with disease characteristics in systemic light chain (AL) amyloidosis patients at diagnosis. Newly diagnosed AL patients were recruited at 2 centers (N=61). Patients completed the PROMIS Global Health v1.2, PROMIS-29 Profile v2.0, and Fatigue 8a v1.0. We assigned disease severity based on stage, presence of cardiac AL, and number of organs involved. We evaluated a) known groups validity by comparing PROMIS T-scores by disease severity, b) internal consistency using Cronbach's alpha, and c) convergent/discriminant validity based on correlations across the domains and summary scores. Using receiver operating characteristic (ROC) curve analysis, NT-proBNP cutoff level corresponding to normal/mild vs moderate/severe PRO scores was determined. The median age was 68 (48–83) years with 58% males. 66% had cardiac involvement and 25% had 3 or more organs involved with AL amyloidosis; 14% had stage 1, 28% stage 2, 36% stage 3 and 16% stage 4 disease. PROMIS measures had acceptable to excellent internal consistency and expected patterns of correlations. PROMIS Global Physical Health score was worse than the Global Mental Health Score at diagnosis; Physical function, fatigue, and anxiety were the most impaired domains. PROMIS Global Health summary scores discriminated across AL amyloidosis stage and number of organs involved. Physical Function showed the strongest effects across known groups by stage, cardiac involvement and number of organs involved followed by Ability to Participate in Social Roles and Activities. A diagnostic NT-proBNP cut-off of 4,200 pg/ml identified patients with moderate/severe PRO scores for these domains. Our results provide evidence for reliability and validity of select PROMIS short form measures in AL amyloidosis at diagnosis.

Corresponding author: Anita D'Souza, MD, MS, Associate Professor of Medicine, Medical College of Wisconsin, Milwaukee, WI 53226, andsouza@mcw.edu, Phone: 414-805-0637, Fax: 414-805-0714.

Disclosure statement

The authors report no conflicts of interest.

Keywords

AL amyloidosis; PROMIS; SF-36; patient-reported outcomes; PROs

Introduction

Light chain (AL) amyloidosis is a systemic disease associated with a clonal plasma cell proliferation in the bone marrow, affecting 10–14 per million person-years in the US [1]. This is a disease of high morbidity and mortality, with a 2 year overall survival of 60% from diagnosis [2]. Symptoms of systemic AL amyloidosis depend on the organs involved with AL deposition, with the heart and kidneys being the most commonly involved organs [3]. Key symptoms include fatigue, shortness of breath, swelling, and pain [4].

Health-related quality of life (HRQoL) in systemic AL amyloidosis has been studied primarily using the Medical Outcomes Study Short Form-36 (SF-36) measure,[5] a commonly-used patient-reported outcome (PRO) measure covering physical and mental health function [6]. Patients with AL amyloidosis have significantly worse HRQoL across all eight SF-36 sub-scales and the two summary scores (physical component summary and mental component summary) compared to the general population with General Health and Role Physical (role limitations due to physical health functioning) being the most affected sub-scales. The content validity and psychometric properties of the SF-36 have been widely studied and it has demonstrated adequate reliability, validity, and sensitivity to change [7, 8] in AL amyloidosis, suggesting that it is an acceptable tool to measure HRQoL in AL amyloidosis patients.

The Patient-Reported Outcomes Measurement Information System® (PROMIS®) is an NIH Roadmap initiative that has advanced the use of a common set of PRO tools to measure patient-reported symptoms, functioning, and other aspects of HRQoL in the general population and across a wide variety of diseases and conditions [9]. The PROMIS network developed item banks and short forms in multiple health domains for adults and children as well as a set of global health items and profile measures. PROMIS development and validation included state-of-the-science qualitative and psychometric methods. PROMIS measures and their scoring are publicly available. Multiple PROMIS scores can be compared to other measures of similar concepts, including the SF-36 and Functional Assessment of Cancer Therapy (FACT) measures [10–17]. When particular domains are measured, scores can also be combined for use in economic evaluations and cost-effectiveness analyses [18]. Patient-reported outcome measures may be able to meet the FDA standards for approval as Clinical Outcomes Assessment (COA) tools, and some PROMIS measures (e.g. Physical Function, Fatigue) have been included as qualified measures in clinical trials testing novel therapies [19]. Thus, we sought to study PROs in AL amyloidosis using PROMIS and herein describe our early findings on the measurement properties of PROMIS Global Health and PROMIS-29 profile in AL amyloidosis, with recommendations for using PROMIS in this context. We hypothesized that patients with more advanced disease as measured by stage, presence of cardiac involvement, and 3 or more organ AL involvement would have worse scores on domains tested. Based on the published literature using the SF-36,[5] we

hypothesized that physical function and fatigue would be the greatest affected PRO domains in AL amyloidosis.

Patients and Methods

We conducted this IRB-approved, prospective cohort study at the Medical College of Wisconsin (N=35) and Mayo Clinic, Rochester (N=26). Newly diagnosed patients with systemic AL amyloidosis within 3 months of starting treatment were eligible for enrollment. In this analysis, we have studied baseline PROs at enrollment. Patients were staged using the Mayo 2004[20] and 2012[21] AL amyloidosis staging systems using NT-proBNP, troponin T, and the difference between the involved and uninvolved free light chains.

PRO measures:

PROMIS Global Health v1.2 is a 10-item scale covering overall evaluations of physical, mental, and social health, with summary scores for Global Physical Health and Global Mental Health scores [22]. In addition, the individual items can be examined separately to provide specific information about perceptions of physical function, pain, fatigue, emotional distress, social health, and general perceptions of health [22, 23].

PROMIS-29 Profile v2.0 is a collection of seven 4-item short forms assessing Anxiety, Depression, Fatigue, Pain Interference, Physical Function, Sleep Disturbance, and Ability to Participate in Social Roles and Activities as well as a single Pain Intensity item. Each domain is scored separately. Domain scores from the PROMIS-29 can be compared to other fixed-length PROMIS short forms as well as scores obtained using computerized adaptive tests [24, 25].

PROMIS Short Form v1.0 – Fatigue 8a is a collection of 8 items (4 items in addition to the 4-item Fatigue domain of the PROMIS-29 profile). This short form asks about the intensity of fatigue and its impact on day-to-day function [23].

We used the HealthMeasures Scoring Service to calculate response pattern-based scores based on item response theory (IRT) for each domain [26]. The IRT scoring approach is advantageous as it takes into account information about each item (e.g., difficulty and discrimination) in addition to the item responses themselves, thus, each response pattern is typically associated with a unique score estimate, allowing for finer-grained measurement than is possible with summed scores [27]. PROMIS scores are represented on the T-score metric (mean = 50, standard deviation (SD) = 10), scaled so that a score of 50 corresponds to the mean of the reference population; for most PROMIS domains the reference population is the general U.S. adult population, though for 2 domains in the PROMIS-29, Ability to Participate in Social Roles and Activities and Sleep Disturbance, the calibration sample included more people with chronic illness. Higher scores indicate more of the concept being measured. For example, for Physical Function domain, a score higher than 50 implies better physical function and score lower than 50 implies worse physical function, whereas for Fatigue, a score higher than 50 implies greater fatigue and lower than 50 means less fatigue compared to the reference (i.e. general US adult population). Minimal important differences vary by domain and the method used to calculate them; with PROMIS measures they may be

as low as 2.0–3.0 points (pain, physical function) [10] or 3.0–5.0 points (fatigue) [28]. For this analysis, we took a conservative approach and considered a difference of half a standard deviation (5 points on T-score) as meaningful.

Statistical analysis: We described domains as mean T-scores with SD. Global Health Summary scores and scores on individual domains were examined to assess differences between known groups by stage, cardiac AL involvement, and number of organs involved. For all inferential analyses, 2012 stages 1 and 2 were combined into a single group, and stages 3 and 4 were combined into a single group due to the small sample size. One-way analysis of variance (ANOVA) with *post-hoc* pairwise comparisons and independent samples *t*-tests were used to test for statistically significant differences across groups on the PRO measures. Due to the exploratory nature of the study, we interpreted all results with *p*-value < 0.1; however, *p* < 0.05 was considered statistically significant. As a measure of effect size, Cohen's *d* was calculated for all pairwise comparisons, regardless of statistical significance (*d* = 0.2 – weak, *d* = 0.5 - medium and *d* = 0.8 – large effect size) [29]. In PROMIS, effect sizes corresponding to minimal important differences have averaged between 0.4 and 0.63 [28]. Because of the small and unequal sample sizes across groups, variances were not assumed to be equal for any of the statistical tests. For the PRO domains that showed a significant relationship with AL disease severity (stage, cardiac involvement, 3 or greater organs involved), we performed a receiver operating characteristic (ROC) curve analysis to determine the NT-proBNP cutoff level at which the best dichotomous PROMIS classification was achieved, with PRO domains dichotomized at 1 SD worse than the reference population mean. This clinically meaningful PROMIS cut-point corresponds to normal/mild versus moderate/severe symptoms. Internal consistency was assessed using coefficient alpha for each domain of the measure, with an alpha > 0.7 considered an acceptable level of internal consistency when using the scales to make group comparisons and > 0.9 (excellent), suitable for individual patient assessments [30]. We evaluated the relationships between domains (convergent/discriminant validity) using Pearson correlation coefficients. Correlations were considered weak if 0.1–0.29, medium if 0.3–0.49 and strong at 0.5 [29]. All statistical analyses were conducted in R version 3.5.2.

Results

We enrolled 61 patients between 2/1/2016 to 4/15/2019. One patient was ineligible based on inclusion criteria (the diagnosis of AL amyloidosis was not confirmed) and one other patient had localized instead of systemic amyloidosis. Among the remaining 59 patients that continued on the study, baseline characteristics are shown in Table 1. Baseline PROMIS scale scores and their distributions are shown in Table 2 and Supplementary Figure, respectively. Modest floor effects (clustering at the best possible score) are noted for a few domains (Pain Interference, Depression, Anxiety, Fatigue 4a, Physical Function, and Ability to Participate in Social Roles and Activities). For the entire population, the mean Global Physical Health Score was low at 42.5 compared to the general population, but the Global Mental Health Score was only slightly lower than the general population at 48.5. The domains with the worst scores (>5 T-score difference from US average) were Anxiety (55.5), Fatigue (55.6), and Physical Function (39.8).

Patient-reported outcome domains most able to differentiate between AL amyloid disease severity groups (Known groups validity):

Table 3 shows the mean T-scores with SD for PRO domains and summary scores by 2012 AL stage, presence of cardiac involvement and number of organs involved with AL.

Disease stage: Stratifying by the 2012 staging system, there was a medium but non-significant effect of stage on Global Physical Health scores (Stages 1–2: Mean = 45.9 [SD = 12.5] vs. Stages 3–4: Mean = 39.5 [SD 11.3], $p = 0.05$, Cohen's $d = 0.54$). There were statistically significant differences in Global Mental Health scores (Stages 1–2: Mean = 51.6 [SD = 10.1] vs. Stages 3–4: M = 45.9 [SD = 8.0], $p = 0.02$, Cohen's $d = 0.63$). The differences in scores by stage were statistically significant and medium in magnitude. Regarding the PROMIS-29 profile, there were significant differences in Physical Function domain (Stages 1–2: Mean = 42.6 [SD = 10.5] vs. Stages 3–4: Mean = 36.8. [SD = 10.4], $p = 0.04$, Cohen's $d = 0.56$), and in Ability to Participate in Social Roles (Stages 1–2: Mean = 51.5 [SD = 9.5], vs. Stages 3–4: Mean = 43.7 [SD = 10.9], $p < .001$, Cohen's $d = 0.76$). Supplementary Figure 2 shows the differences across all PRO domains by 2012 staging.

Cardiac AL involvement: The only domain that differentiated between patients with cardiac AL and those without was Physical Function (with cardiac AL: Mean = 37.7 [SD = 10.6] vs. without cardiac AL: Mean = 44 [SD = 10.3], $p = 0.03$, Cohen's $d = 0.6$). Though not statistically significant, a medium effect size was also observed for Sleep Disturbance (with cardiac AL: Mean = 53.57 [SD = 9.39] vs without cardiac AL: Mean = 48.37 [SD 10.23], $p = 0.07$, Cohen's $d = 0.53$).

Number of organs involved with amyloid: In comparing across groups based on the number of organs involved, large differences were found for PROMIS Global Physical Health (3/+ organs: Mean = 35.0 [SD 9.7] vs. 1 organ: Mean = 45.8 [SD 12.1], $p < 0.01$, Cohen's $d = 0.85$ and vs. 2 organs: M = 44.21 [SD = 11.81], $p = 0.01$, $d = 0.85$), and Global Mental Health (3/+ organs: Mean = 42.2 [7.5] vs. 1 organ: Mean = 50.1 [SD = 8.8], $p < 0.01$, Cohen's $d = 0.98$ and vs. 2 organs: M. = 51.0 [SD = 9.44], $p < 0.01$, Cohen's $d = 1.03$). On the PROMIS-29 profile, while the overall test for Fatigue was not statistically significant ($p = .08$), there were significant pairwise differences between specific groups (3/+ organs: Mean = 61.3 [SD 11.2] vs. 1 organ: Mean = 52.3 [SD 11.8], $p = 0.03$, Cohen's $d = 0.78$). Physical Function showed a significant overall test as well as a large effect size for pairwise comparisons (3/+ organs: Mean = 34.4 [SD = 10.6] vs. 1 organ: Mean = 43.6 [SD = 9], $p = 0.01$, Cohen's $d = 0.92$), as did Ability to Participate in Social Roles (3/+ organs: Mean = 39.9 [SD= 11.2] vs. 1 organ: Mean= 49.7 [SD = 9.7], $p = 0.01$, Cohen's $d = 0.94$).

ROC curve analysis:

Based on these analyses, we concluded that physical function and social roles participation were the most affected PRO domains in AL amyloidosis with increasing disease severity. For these domains, we conducted an ROC curve analysis to identify a cut-point in NT-proBNP that most accurately classifies patients into normal/mild vs. moderate/severe PROMIS scores for these domains. An NT-proBNP cut-point of 4,200 pg/ml was able to discriminate between normal/mild (>40) and moderate/severe (< 40) Physical Function

scores with a sensitivity of 55%, specificity of 79%, and area under the curve of 0.64 and Ability to Participate in Social Roles and Activities with a sensitivity of 79%, specificity of 72% and area under the curve of 0.77 (Figure 1).

Internal consistency of PROMIS in AL amyloidosis:

Coefficient alpha was high across all individual domains, with 0.94 for Physical Function, 0.92 for Anxiety, 0.90 for Depression, 0.97 for Fatigue, 0.89 for Sleep Disturbance, 0.97 for Ability to Participate in Social Roles and Activities, and 0.98 for Pain. It was somewhat lower, though still acceptable, for PROMIS Global Health Summary scores, with 0.84 for Global Physical Health and 0.78 for Global Mental Health.

Convergent/Discriminant validity:

We assessed correlations among individual PROMIS domains and the PROMIS Global Health scales, as well as among the items within the Global Health scale and its summary scores (Table 4). Patterns were mostly as expected (based on patterns observed in the general population) [22], with higher correlations among physical health domains (Physical Function, Pain, and Fatigue, correlations ranging 0.42–0.77) and among mental health domains (Anxiety and Depression, correlations ranging 0.64–0.79) than across physical and mental health domains (correlations ranging 0.34–0.63). Sleep Disturbance had weak to medium, but significant, correlations with all other domains in the PROMIS-29 (0.29–0.40). Within the Global Health scale, the items representing physical function, physical activities, pain, and fatigue were more highly correlated with the Global Physical Health score, however the item representing pain had only moderate correlation with physical function, physical activities, and fatigue (0.35–0.46). Similarly, the items representing quality of life, mental health, satisfaction with social activities, and emotional problems were highly correlated with the Global Mental Health score, but the item representing emotional problems had moderate correlations with overall quality of life, mental health, and satisfaction with social activities (0.31–0.50).

Discussion

Findings from this study provide insight into PROs in AL amyloidosis at the time of diagnosis and into PROMIS PROs specifically. As hypothesized, we found that physical health was significantly impaired at diagnosis. Though the Global Mental Health mean score was also lower than the US average, it was not as low as the Global Physical Health score. Physical function, fatigue, and anxiety were also highly impaired in this group at diagnosis. PRO scores were significantly different, or exhibited moderate to large effect size differences by AL stage, number of AL organs involved, and cardiac AL involvement. The largest impacts at diagnosis were related to physical functioning and social role participation. Our data also identified that an NT-proBNP cut-point of 4,200 pg/ml as able to discriminate between no/mild and moderate/severe scores for physical function and social roles. PROMIS measures had acceptable to excellent internal consistency, and patterns of correlations were mostly as expected. We tested both the 4- and 8-item short forms for Fatigue. While choice of a shorter or longer measure is dependent on specific context and objectives, based on the score distributions and correlations with other domains, either the 4-

or 8-item version is likely appropriate in this population. As PROMIS measures are increasingly being used across multiple research and clinical care settings, it is important to understand how they perform in specific disease contexts. This study evaluating PROMIS measures in AL amyloidosis adds to our previous work on PROMIS [31] and recent retrospective work by others [32]. Our study is also one of the first in understanding PRO domains most affected at diagnosis in AL amyloidosis.

From previous work using other measures, patients with AL have been shown to have broad HRQoL deficits relative to the general population, with the largest effects seen in physical functioning and general well-being [5]. Fatigue is one of the important symptoms reported by amyloid patients [4] and other important symptoms include limitations on physical activities, ability to carry out particular roles and emotional wellbeing [8]. Further, sleep disturbances might be secondary to anxiety, sleep apnea (or other cardiac-related sleep disorders) and pain. The PROMIS measures we used were able to measure anxiety, fatigue, and sleep disturbance scores - scores that are not derived by the SF-36. While the SF-36 measures vitality, one could argue that that is not the same concept as fatigue thus underscoring the need to separately measure fatigue. In a cross-sectional, community-based sample of AL patients with analysis of baseline data using the SF-36v2 measure, greater impairments were seen in patients with recent diagnosis and those with cardiac involvement [5]. This is similar to our data which was limited to newly diagnosed AL patients and showed that the Global Physical Health summary score, along with physical functioning, fatigue, and social roles participation were the most significantly affected. The Ability to Participate in Social Roles and Activities domain includes questions identifying satisfaction with how much work one can do, ability to work, ability to do regular personal and household responsibilities and to perform daily routines. In AL patients, this domain may be reflective of poor physical function or could be indicative of time spent dealing with illness. Further qualitative work may clarify this difference.

While anxiety was high overall, we found no differences in anxiety or depression by AL disease markers, yet the PROMIS Global Mental Health score was significantly different by AL stage. In past work, nearly 37% of AL patients report depression defined as at least 1 depressive symptom for at least “a good bit of the time” during the 4 weeks prior to completing the SF-36, and 47% report anxiety defined as endorsement of at least 1 anxiety symptom for at least “a good bit of the time” [33]. However, when assessing the SF-36 study by cardiac AL involvement, mental health was not different by cardiac involvement, with the Role Emotional domain mean score of 45.3 (SD 12.5) with versus 45.3 (SD 12) without ($p=0.9$), Mental Health domain mean score 49.2 (SD 10.5) with and 48.9 (SD 10.8) without ($p=0.8$) and Mental Component Summary mean score of 48.5 (SD 11.4) with and 47.2 (SD 11.9) without ($p=0.4$) cardiac AL [5]. In another study of distress in patients with systemic AL amyloidosis using the National Comprehensive Cancer Network Distress Thermometer scale, patient-reported distress was not associated with stage or organ involvement, and patients with cardiac AL reported lower distress compared to patients without cardiac involvement ($p=0.02$) [34]. The PROMIS Global Mental Health score includes 4 questions: rate overall quality of life, how often have you been bothered by emotional problems such as feeling anxious, depressed or irritable, how would you rate your mental health, including your mood and your ability to think, and how would you rate your social activities and

relationships. Thus, it is possible that the Global Mental Health score difference we observed was driven by the question on social activities and relationships, which is similar to the Ability to Participate in Social Roles and Relationships domain that was significant across multiple analyses in our study. We have previously hypothesized that the discordance between cardiac AL patients not experiencing worse anxiety/depression may be secondary to relief patients feel after receiving a definitive diagnosis [34]. Delayed diagnosis is common in amyloidosis, with over 38% of patients having a lag of over a year between initial symptoms to diagnosis and nearly half of patients needing to see four or more different physicians before the diagnosis is established [35]. Qualitative research to understand this aspect will be helpful in understanding this discordance.

The NT-proBNP has been shown to be highly prognostic for AL amyloidosis outcomes, including for worsening stage (2004 Mayo IIIb) at a cut-point of 8,500 pg/ml [36], risk of early mortality at a cut-point of 4,200 pg/ml [37], and risk of transplant-related mortality following an autologous stem cell transplant at a cut-point of 5,000 pg/ml [38]. Similar to these reports, an NT-proBNP cut-off of 4,200 pg/ml at diagnosis was able to predict worse patient-reported function across the Physical Function and Ability to Participate in Social Roles and Activities domains.

The internal consistency of the instruments was similar to what was demonstrated in the general population with excellent internal consistency for domains within the PROMIS-29 and acceptable internal consistency for the Global Health summary scores [22, 39]. The Global Health measure was designed for population-level health monitoring and is less appropriate for individual patient assessment. That said, the internal consistency is adequate for group comparisons, even in small samples, such as those found in rare disease contexts like AL.

Our study is limited by the small sample size and thus limited power for statistical inference. Within our known groups of disease severity, Fatigue showed a large score difference (of >5 points on the T-score metric) between groups but did not meet our statistical threshold for significance and this may well be owing to our small sample size. We did not include other measures such as KCCQ or FACIT to compare the convergent/discriminant validity of PROMIS in AL amyloidosis, but this is an area for future research. Despite the small sample size in our study, we are encouraged that we were able to make substantive additions to the understanding of HRQoL measurement in AL disease, identify important PRO domains that need to be studied, and the next steps needed to advance this work. Our next steps include qualitative research to further understand the mental health domains in AL amyloidosis, assessing changes in PROs over time, and expanding the work to larger groups of patients. Ultimately, the goal of our work is to identify PROs that may be most useful in determining change in status and could serve as a clinical outcomes assessment tool in AL clinical trials.

In summary, our study identifies physical function, fatigue, and anxiety as the most impaired domains at diagnosis in AL amyloidosis. Physical function and social role participation have the most ability to differentiate by disease severity at diagnosis. An NT-proBNP cut-off of 4,200 pg/ml can discriminate a change in 1 SD in these domains. Our study provides initial evidence of psychometric properties such as reliability and validity for PROMIS Global

Health and PROMIS-29 measures in AL amyloidosis. Additional qualitative work needs to be done to better understand and measure mental health domains in AL amyloidosis patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors gratefully acknowledge assistance from the Clinical Research Coordinators at Mayo Clinic and MCW in enrolling patients, data collection, and maintenance. Research reported in this publication was supported by K23 HL141445, the Research and Education Program of Advancing Healthier Wisconsin and the Clinical and Translational Science Institute (CTSI) (KL2 TR001438). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abbreviations

AL amyloidosis	light chain amyloidosis
HRQoL	health-related quality of life
PRO	patient-reported outcome
PROMIS	Patient-Reported Outcomes Measurement Information System
SF-36	Medical Outcomes Study Short Form-36

References

- [1]. Quock TP, Yan T, Chang E, et al. Epidemiology of AL amyloidosis: a real-world study using US claims data. *Blood Adv.* 2018;2:1046–1053. [PubMed: 29748430]
- [2]. Quock TP, Chang E, Munday JS, et al. Mortality and healthcare costs in Medicare beneficiaries with AL amyloidosis. *J Comp Eff Res.* 2018;7:1053–1062. [PubMed: 30354284]
- [3]. Merlini G AL amyloidosis: from molecular mechanisms to targeted therapies. *Hematology Am Soc Hematol Educ Program* 2017;2017:1–12. [PubMed: 29222231]
- [4]. Lin HM, Seldin D, Hui AM, et al. The patient's perspective on the symptom and everyday life impact of AL amyloidosis. *Amyloid.* 2015;22:244–251. [PubMed: 26587688]
- [5]. Bayliss M, McCausland KL, Guthrie SD, et al. The burden of amyloid light chain amyloidosis on health-related quality of life. *Orphanet J Rare Dis.* 2017;12:15. [PubMed: 28103898]
- [6]. Maruish ME. User's manual for the SF-36v2 Health Survey. Quality Metric Incorporated, 2011.
- [7]. White MK, McCausland KL, Sanchowala V, et al. Psychometric validation of the SF-36 Health Survey in light chain amyloidosis: results from community-based and clinic-based samples. *Patient Relat Outcome Meas.* 2017;8:157–167. [PubMed: 29263707]
- [8]. White MK, Bayliss MS, Guthrie SD, et al. Content validation of the SF-36v2(R) health survey with AL amyloidosis patients. *J Patient Rep Outcomes.* 2017;1:13. [PubMed: 29757308]
- [9]. Rothrock NE, Hays RD, Spritzer K, et al. Relative to the general US population, chronic diseases are associated with poorer health-related quality of life as measured by the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol.* 2010;63:1195–1204.
- [10]. Chen CX, Kroenke K, Stump T, et al. Comparative responsiveness of the PROMIS Pain Interference Short Forms With Legacy Pain Measures: results from three randomized clinical trials. *J Pain.* 2019;20:664–675. [PubMed: 30529442]

- [11]. Kaat AJ, Schalet BD, Rutsohn J, et al. Physical function metric over measure: An illustration with the Patient-Reported Outcomes Measurement Information System (PROMIS) and the Functional Assessment of Cancer Therapy (FACT). *Cancer*. 2018;124:153–160. [PubMed: 28885707]
- [12]. Schalet BD, Revicki DA, Cook KF, et al. Establishing a common metric for physical function: linking the HAQ-DI and SF-36 PF Subscale to PROMIS((R)) Physical Function. *J Gen Intern Med*. 2015;30:1517–1523. [PubMed: 25990189]
- [13]. Cook KF, Schalet BD, Kallen MA, et al. Establishing a common metric for self-reported pain: linking BPI Pain Interference and SF-36 Bodily Pain Subscale scores to the PROMIS Pain Interference metric. *Qual Life Res*. 2015;24:2305–2318. [PubMed: 25894063]
- [14]. Choi SW, Schalet B, Cook KF, et al. Establishing a common metric for depressive symptoms: linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. *Psychol Assess*. 2014;26:513–527. [PubMed: 24548149]
- [15]. Schalet BD, Cook KF, Choi SW, et al. Establishing a common metric for self-reported anxiety: linking the MASQ, PANAS, and GAD-7 to PROMIS Anxiety. *J Anxiety Disord*. 2014;28:88–96. [PubMed: 24508596]
- [16]. Lai JS, Cella D, Yanez B, et al. Linking fatigue measures on a common reporting metric. *J Pain Symptom Manage*. 2014;48:639–648. [PubMed: 24698661]
- [17]. PROsetta Stone Linking Patient-Reported Outcome Measures. In. <http://www.prosetastone.org>.
- [18]. Dewitt B, Feeny D, Fischhoff B, et al. Estimation of a Preference-Based Summary Score for the Patient-Reported Outcomes Measurement Information System: The PROMIS((R))-Preference (PROPr) Scoring System. *Med Decis Making*. 2018;38:683–698. [PubMed: 29944456]
- [19]. <https://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm625989.htm>. In. 2018.
- [20]. Dispenzieri A, Gertz MA, Kyle RA, et al. Serum cardiac troponins and N-terminal pro-brain natriuretic peptide: a staging system for primary systemic amyloidosis. *J Clin Oncol*. 2004;22:3751–3757. [PubMed: 15365071]
- [21]. Kumar S, Dispenzieri A, Lacy MQ, et al. Revised prognostic staging system for light chain amyloidosis incorporating cardiac biomarkers and serum free light chain measurements. *J Clin Oncol*. 2012;30:989–995. [PubMed: 22331953]
- [22]. Hays RD, Bjorner JB, Revicki DA, et al. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Qual Life Res*. 2009;18:873–880. [PubMed: 19543809]
- [23]. <http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/list-of-adult-measures>. In. 2018.
- [24]. Choi SW, Reise SP, Pilkonis PA, et al. Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. *Qual Life Res*. 2010;19:125–136. [PubMed: 19941077]
- [25]. Flynn KE, Dew MA, Lin L, et al. Reliability and construct validity of PROMIS(R) measures for patients with heart failure who undergo heart transplant. *Qual Life Res*. 2015;24:2591–2599. [PubMed: 26038213]
- [26]. HealthMeasures Scoring Instructions. <http://www.healthmeasures.net/score-and-interpret/calculate-scores/scoring-instructions>
- [27]. Thissen D, Pommerich M, Billeaud K, et al. Item response theory for scores on tests including polytomous items with ordered responses. *Appl Psychol Measurm*. 1995;19:39–49.
- [28]. Yost KJ, Eton DT, Garcia SF, et al. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System–Cancer scales in advanced-stage cancer patients. *J Clin Epidemiol*. 2011;64:507–516. [PubMed: 21447427]
- [29]. Cohen J *Statistical Power Analysis for the Behavioural Sciences* (2nd Edition). Hillsdale, NJ: Lawrence Erlbaum Associates,1988.
- [30]. Fayers PM, Machin D. *Quality of Life*. Wiley Blackwell,2016.
- [31]. D'Souza A, Hari P, Pasquini M, et al. Baseline patient-reported outcomes in light-chain amyloidosis patients enrolled on an interventional clinical trial. *Amyloid*. 2019;26:87–88. [PubMed: 31343353]

- [32]. Chakraborty R, Rybicki L, Tomer J, et al. Patient-reported outcomes in systemic AL amyloidosis with functional assessment of cancer therapy-general (FACT-G) and patient-reported outcomes measurement information system-global health (PROMIS-GH) in a real-world population. *Leuk Lymphoma*. 2019;60(14):3544–3551. [PubMed: 31272258]
- [33]. Shu J, Lo S, Phillips M, et al. Depression and anxiety in patients with AL amyloidosis as assessed by the SF-36 questionnaire: experience in 1226 patients. *Amyloid*. 2016;23:188–193. [PubMed: 27460276]
- [34]. Wright NL, Flynn KE, Brazauskas R, et al. Patient-reported distress is prevalent in systemic light chain (AL) amyloidosis but not determined by severity of disease. *Amyloid*. 2018; 25: 129–134. [PubMed: 30032653]
- [35]. Lousada I, Comenzo RL, Landau H, et al. Light chain amyloidosis: patient experience survey from the Amyloidosis Research Consortium. *Adv Ther*. 2015; 32: 920–928. [PubMed: 26498944]
- [36]. Wechalekar AD, Schonland SO, Kastiris E, et al. A European collaborative study of treatment outcomes in 346 patients with cardiac stage III AL amyloidosis. *Blood*. 2013;121:3420–3427. [PubMed: 23479568]
- [37]. Kumar SK, Gertz MA, Lacy MQ, et al. Recent improvements in survival in primary systemic amyloidosis and the importance of an early mortality risk score. *Mayo Clin Proc*. 2011;86:12–18. [PubMed: 21193650]
- [38]. Gertz MA, Lacy MQ, Dispenzieri A, et al. Refinement in patient selection to reduce treatment-related mortality from autologous stem cell transplantation in amyloidosis. *Bone Marrow Transplant*. 2013;48:557–561. [PubMed: 22964596]
- [39]. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol*. 2010;63:1179–1194. [PubMed: 20685078]

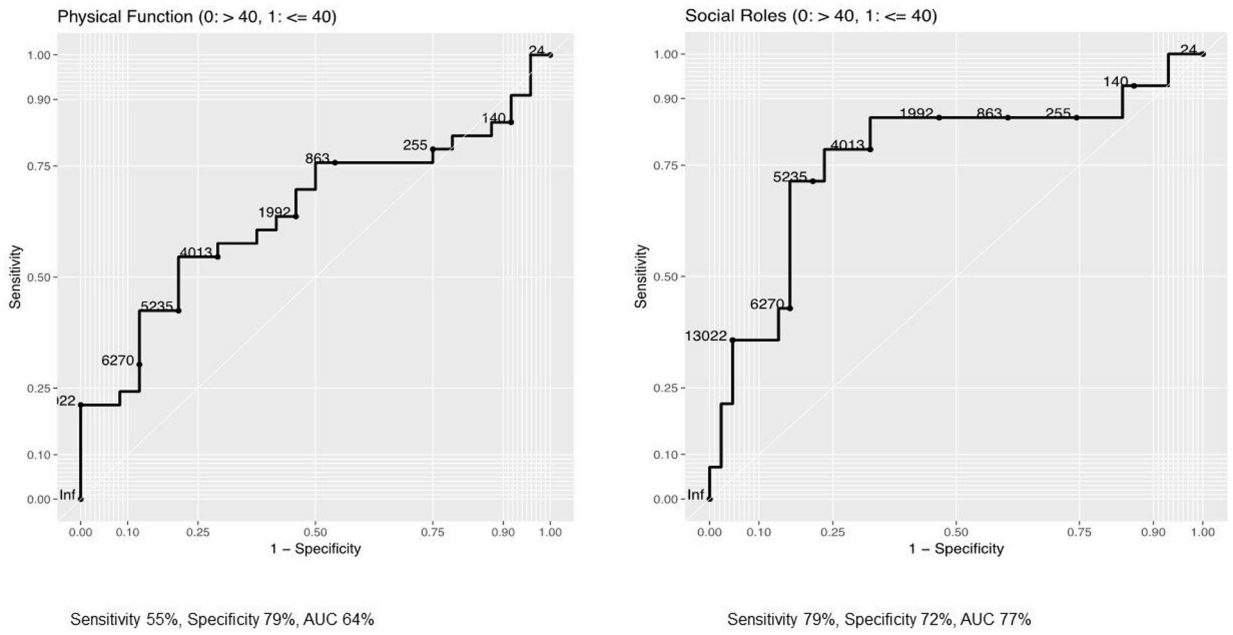


Figure 1.
PROMIS scores across disease severity groups

Table 1.

Baseline characteristics

Characteristic	N=59
Age, median (range)	68 (48–83)
Gender (%)	
-Male	34 (58)
-Female	25 (42)
Race (%)	
-White	53 (90)
-Black	5 (8)
-Others	1 (2)
AL type (%)	
-Kappa	13 (22)
-Lambda	46 (78)
2004 stage (%)	
I	14 (24)
II	17 (29)
III	26 (44)
Missing *	2 (3)
2012 stage (%)	
1	9 (14)
2	18 (28)
3	20 (36)
4	10 (16)
Missing *	2 (6)
Median (range) dFLC, mg/L	82 (2–978)
Median (range) NT proBNP, pg/mL	2,643 (24–50,863)
Median (range) troponin T, ng/mL	0.05 (<0.01–206)
Cardiac AL (%)	39 (66)
Renal AL (%)	34 (58)
Number of AL organs involved	
1	21 (36%)
2	23 (39%)
3/>	15 (25%)

* 3 patients had missing biomarkers prior to starting therapy

dFLC- difference between involved and uninvolved free light chain

Table 2.

Baseline PROMIS scores

Measure	Baseline Mean (SD)
PROMIS Global Health	
-Physical Health Summary Score	42.5 (12.1)
-Mental Health Summary Score	48.5 (9.4)
Individual domains	
-Anxiety	55.5 (8.7)
-Depression	53.4 (9.2)
-Fatigue	55.6 (12.2)
-Pain Interference	51.2 (10.9)
-Physical Function	39.8 (10.8)
-Sleep disturbance	51.8 (9.9)
-Ability to Participate in Social roles and Activities	47.1 (10.9)

Score interpretation: Higher scores indicate more of the concept being measured. For the domains of Anxiety, Depression, Fatigue (derived from Fatigue 8a), Pain, Sleep Disturbance, a higher score represents greater symptom; Physical Function, Ability to Participate in Social Roles and Activities, Global Physical Health, Global Mental Health: higher score represents better function

PROMIS scores are represented on the T-score metric (mean = 50, standard deviation (SD) = 10), scaled so that a score of 50 corresponds to the mean of the reference population; for most PROMIS domains the reference population is the general U.S. adult population, though for 2 domains in the PROMIS-29, Ability to Participate in Social Roles and Activities and Sleep Disturbance, the calibration sample included more people with chronic illness.

Table 3.

PROMIS domain scores across disease severity spectrum

	Stage 1/2 Mean (SD)	Stage 3/4 Mean (SD)	p-value	Cohen's <i>d</i>
Pain Interference	50.5 (11.1)	51.2 (11.0)	0.8	0.07
Depression	52.3 (9.5)	54.1 (9.2)	0.5	0.20
Physical Function **	42.6 (10.5)	36.8 (10.4)	0.04	0.56
Social Roles ***	51.5 (9.5)	43.7 (10.9)	<0.001	0.76
Fatigue	53.2 (12.2)	58.0 (11.9)	0.1	0.40
Anxiety	54.8 (8.9)	55.9 (8.9)	0.6	0.13
Sleep Disturbance	49.4 (8.5)	53.6 (10.8)	0.1	0.43
Global Physical Health **	45.9 (12.5)	39.5 (11.3)	0.05	0.54
Global Mental Health **	51.6 (10.1)	45.9 (8.0)	0.02	0.63

	No Cardiac AL Mean (SD)	Cardiac AL Mean (SD)	p-value	Cohen's <i>d</i>
Pain Interference	54.5 (11.5)	49.6 (10.3)	0.1	0.44
Depression	51.7 (8.6)	54.3 (9.5)	0.3	0.30
Physical Function **	44.0 (10.3)	37.7 (10.6)	0.03	0.60
Social Roles	49.0 (11.2)	46.0 (10.7)	0.3	0.27
Fatigue	52.6 (13.0)	57.1 (11.6)	0.2	0.36
Anxiety	54.4 (7.5)	56.1 (9.4)	0.5	0.19
Sleep Disturbance *	48.4 (10.2)	53.6 (9.4)	0.07	0.53
Global Physical Health	45.1 (12.7)	41.1 (11.7)	0.3	0.32
Global Mental Health	50.2 (10.5)	47.6 (8.7)	0.3	0.27

	1 Organ Mean (SD)	2 Organs Mean (SD)	3+ Organs Mean (SD)	Overall p-value	1 vs. 3 <i>p</i> (Cohen's <i>d</i>)	2 vs. 3 <i>p</i> (Cohen's <i>d</i>)
Pain Interference *	51.3 (10.0)	47.7 (8.7)	56.8 (13.4)	0.08	0.1 (0.47)	0.01 (0.80)
Depression	53.6 (9.0)	52.0 (9.3)	55.4 (9.5)	0.6	0.6 (0.20)	0.3 (0.36)
Physical Function **	43.6 (9.4)	39.8 (11.1)	34.4 (10.6)	0.04	0.01(0.92)	0.1 (0.50)
Social Roles ***	49.7 (9.7)	49.3 (9.9)	39.9 (11.2)	0.02	0.01 (0.94)	0.01 (0.89)
Fatigue *	52.3 (11.8)	54.8 (12.2)	61.3 (11.2)	0.08	0.03 (0.78)	0.1 (0.55)
Anxiety	55.6 (9.1)	53.9 (9.0)	57.8 (7.9)	0.4	0.4 (0.26)	0.2 (0.47)
Sleep Disturbance	49.4 (10.1)	52.1 (9.4)	54.7 (10.3)	0.3	0.1 (0.52)	0.4 (0.26)
Global Physical Health ***	45.8 (12.1)	44.2 (11.8)	35.0 (9.7)	0.01	<0.01(0.98)	0.01(0.85)
Global Mental Health ***	50.1 (8.8)	51.0 (9.4)	42.2 (7.5)	<0.01	<0.01(0.98)	<0.01 (1.03)

Fatigue is derived from Fatigue 8a

Cohen's *d* is a measure of effect size: $d=0.2-0.49$ - weak, $0.5-0.79$ - medium, and >0.8 - large effect sizes

Domains and summary scores are marked significant for exploration:

*
p <0.1,

**
p <0.05,

p <0.01

Bolded domains and summary scores are ones with p-value <0.5 and effect size >0.5

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4.

Correlation among PROMIS scales

Correlations between PROMIS Global Health, PROMIS-29, and Fatigue-8 in AL amyloidosis									
	Global PH	Global MH	Anxiety	Depression	Fatigue	Pain	Physical Function	Sleep Disturbance	Social Roles
Global PH	1.00	0.74	-0.45	-0.60	-0.86	-0.55	0.86	-0.32	0.77
Global MH		1.00	-0.68	-0.73	-0.62	-0.60	0.59	-0.38	0.75
Anxiety			1.00	0.79	0.46	0.45	-0.37	0.40	-0.41
Depression				1.00	0.64	0.40	-0.46	0.40	-0.49
Fatigue					1.00	0.51	-0.77	0.39	-0.72
Pain						1.00	-0.42	0.36	-0.56
Physical Function							1.00	-0.29	0.75
Sleep Disturbance								1.00	-0.35
Social Role									1.00

Correlations within PROMIS Global Health Scale in AL amyloidosis											
	Global 03	Global 06	Global 07	Global 08	Global PH		Global 02	Global 04	Global 05	Global 10	Global MH
Global 03	1.00	0.74	0.46	0.70	0.87	Global 02	1.00	0.62	0.63	0.33	0.83
Global 06		1.00	0.35	0.77	0.87	Global 04		1.00	0.48	0.50	0.82
Global 07			1.00	0.45	0.66	Global 05			1.00	0.31	0.81
Global 08				1.00	0.89	Global 10				1.00	0.67
Global PH					1.00	Global MH					1.00

Fatigue is derived from Fatigue 8a

Interpretation: The numbers in the boxes signify Pearson's correlation coefficients; Score 0.1–0.29- weak, 0.3–0.49- medium and 0.5- strong correlations. Positive numbers signify correlation in the same direction and negative indicate correlation in opposite directions. e.g. the domains of Depression and Fatigue have a Pearson correlation coefficient of 0.64 implying a strong positive relation between the 2 domains, while Fatigue and Social Roles have a Pearson correlation coefficient of -0.72 implying a strong negative correlation between the 2 domains

Global PH- PROMIS Global Physical Health Summary Score (sum of global 03, 06, 07 and 08); Global MH- PROMIS Mental Health Summary Score (sum of global 02, 04, 05, 10).

Global 03- Physical health; Global 06- Physical function; Global 07- Pain; Global 08- Fatigue

Global 02- Quality of life; Global 04- Mental health; Global 05- Satisfaction with social activities; Global 10- Emotional problems