


Stakeholder Development of an Online Program to Track Arthritis-Related Patient-Reported Outcomes Longitudinally: Live Yes! INSIGHTS

Karen E. Schifferdecker,^{1,2}  Rebecca L. Butcher,^{1,2} Erin Knight,^{1,2} Emily Creek,³ M. Suzanne Schrandt,³ Laura Marrow,³ Marie Jaffe,³ Arlene Vinci,³ and Guy Eakin³

Objective. Patient-reported outcome measures (PROMs) are increasingly used in clinical settings but may not provide benefits to patients outside of health encounters. The Arthritis Foundation's Live Yes! Network provides an opportunity for PROM use by individuals and the network that assists individuals with managing their arthritis between encounters. Our objective was to develop a patient-reported outcomes platform for the network, Live Yes! INSIGHTS, using mixed methods and extensive stakeholder input.

Methods. A mixed methods longitudinal transformation design was used, starting with semistructured interviews to specify the main priorities of the program, literature review to identify potential PROMs, modified Delphi and nominal group technique to select final PROMs, and focus groups to guide program implementation, messaging, and use of results. We gathered input from 93 participants, including from individuals living with arthritis (74% of participants), caregivers, providers, researchers, and measurement experts.

Results. Our mixed methods study resulted in the selection of Patient-Reported Outcomes Measurement Information System (PROMIS)-29, PROMIS Emotional Support Short Form v2.0, and the Health Care Empowerment Questionnaire, to be deployed through a Qualtrics platform. Triangulation of data resulted in identification of potential risks and benefits, including confidentiality, ability to personally track and share data, and an opportunity to contribute to research.

Conclusion. An accessible measurement system backed by psychometrically strong PROMs, created with robust stakeholder engagement, and linked to a national patient network sets the stage for individuals with arthritis to better monitor and improve health outcomes both outside and inside health care settings and for the network to customize programming to meet needs.

INTRODUCTION

Calls for the use of patient-reported outcome measures (PROMs), defined as “report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else,” (1) continue to grow (2). PROMs have been used in clinical trials and some registries for a number of years (3–5). Today, their use is growing in clinical settings with measurable impact on patient care, shared decision making (6–8), and management of population health (9,10).

The arthritis community, which includes patients representing more than 100 disease categories, is one of the most active in promoting the use of PROMs (11), with support of a number of specialties including rheumatology, orthopedics, and physical therapy. These communities have promoted PROMs for arthritis in sophisticated ways, including the development and adaptation of PROMs, establishment of recommendations for using PROMs in quality measurement, (12) and research. This has resulted in several disease-specific, arthritis-related PROMs (13–16) and validation of generic PROMs for arthritis-related conditions (12,17–19). PROMs are also included in several rheumatology research and quality

This work was supported through a contract funded by the Arthritis Foundation. The authors report no other financial interests.

¹Karen E. Schifferdecker, PhD, MPH, Rebecca L. Butcher, MS, MPH, Erin Knight, PhD: Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, New Hampshire; ²Karen E. Schifferdecker, PhD, MPH, Rebecca L. Butcher, MS, MPH, Erin Knight, PhD: Dartmouth College Geisel School of Medicine, Hanover, New Hampshire; ³Emily Creek, MBA, M. Suzanne Schrandt, JD, Laura Marrow, Marie Jaffe, Arlene Vinci, Guy Eakin, PhD: Arthritis Foundation, Atlanta, Georgia.

No potential conflicts of interest relevant to this article were reported

Address correspondence to Karen E. Schifferdecker, PhD, MPH, Geisel School of Medicine at Dartmouth, Center for Program Design and Evaluation, WTRB, Level 5, 1 Medical Center Dr., Lebanon, NH 03756. Email: karen.e.schifferdecker@dartmouth.edu.

Submitted for publication June 12, 2020; accepted in revised form October 23, 2020.

registries (eg, CARRA, RISE, PROCOIN) (20,21). In research, the Outcome Measures in Rheumatology (OMERACT) initiative identified tools, including PROMs, that best measure outcomes in rheumatology (22,23).

Although the arthritis community supports use of PROMs, barriers still exist to realizing their full potential in health care settings (24,25). These include difficulty incorporating them into electronic health records, disruption in clinical workflows, and producing real-time results that are useful to providers and patients. In addition, patients who fill out PROMs, whether for a clinical visit or registry, do not consistently know their results or how their results are being used unless shared during an appointment. In response to this, there have been calls for registries to increase patient involvement during and outside of appointments (12,26).

One avenue that can be used to increase use of PROMs by and for patients is to promote their use through patient networks, which can offer value outside of, but complementary to, clinical use. One such network, the Live Yes! Arthritis Network (27), was created by the Arthritis Foundation to provide resources, education, and support for individuals living with arthritis. The network facilitates many community interactions planned by volunteers across the United States, including evidence-based interventions such as Walk With Ease (28), 178 support groups led by 227 trained facilitators, and online communities with close to 14,000 unique users. In addition, the Foundation's website receives approximately 20 million visits annually. Given this network's extensive reach, having access to data on community needs is crucial for setting priorities.

In developing the Live Yes! Network, the Arthritis Foundation collected preliminary data to identify primary drivers of interaction within the network and patient needs and expectations of the network (unpublished). Three focal areas for greater support and resources emerged: better physical health, better emotional and social health, and improved experience of care in interactions with health care providers. Although the Live Yes! Network has a strong core of volunteers, staff, and advisors to provide some information on these areas, it lacked a systematic and standardized mechanism to assess and monitor the three areas across members, regions, and conditions over time. Therefore, the Foundation set

out to develop a mechanism to assist staff and volunteers in tracking member needs in these three domains and to create a data-driven approach for development of responsive programming and strategies.

To this end, our objective was to create a longitudinal PROM measurement platform for the Arthritis Foundation, since named Live Yes! INSIGHTS, or INSIGHTS, to monitor and report on the three focal areas: physical health, emotional and social health, and experience of care. In this paper, we describe the longitudinal mixed methods study we used to create INSIGHTS using established PROMs and extensive stakeholder involvement.

PATIENTS AND METHODS

The overall approach used to develop INSIGHTS was a mixed methods longitudinal transformation design (29). In this approach, qualitative and quantitative data are collected over time, and data are analyzed and triangulated throughout. Based on institutional review board review by Advarra® and Dartmouth's Committee for the Protection of Human Subjects, the study was considered to be exempt from further review per Category 2 regulations.

Figure 1 provides an overview of the timeline and data collection methods for the study. As shown, the study was divided into four major steps: interviews with stakeholders, a critical literature review, a modified nominal group technique (NGT), and focus groups. During the process, we used method and investigator triangulation to assess the trustworthiness of our findings and whether we had reached saturation (30). Below we describe each method and purpose.

Stakeholder interviews. The purpose of interviews was to obtain input on the idea of creating a PROM-based data platform for the Live Yes! network, including if and why the Foundation should create a platform; what kind of data should be gathered; the potential uses of the data for the Foundation, its constituents, and external stakeholders; and considerations related to data sharing. We used purposive sampling to identify a mix of stakeholders (n = 12), including individuals with arthritis (n = 3),

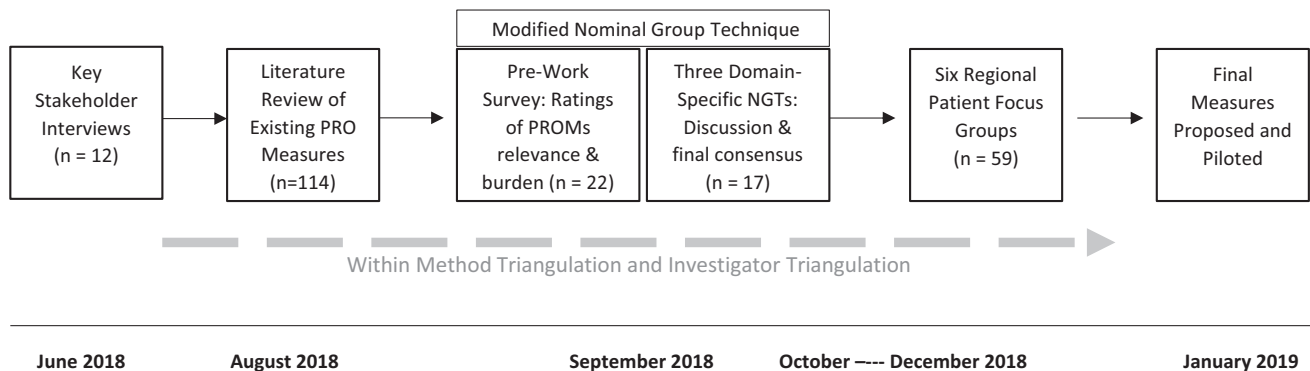


Figure 1. Project timeline and data collection methods

Foundation leaders ($n = 3$), researchers focused on rheumatological conditions ($n = 5$), and industry partners ($n = 1$). We set a target of 12 interviews, as research suggests this is sufficient to reach saturation on a focused topic (31). All interviews were conducted by two non-Foundation academic researchers (RLB, KES) using a semistructured interview guide, audio recorded and analyzed using content analysis in Dedoose, a qualitative data analysis program (32). Top-level findings from the stakeholder interviews validated the three target domains for the platform and informed the critical literature review and subsequent data collection.

Critical literature review. Given additional assurance for 1) the idea of INSIGHTS and 2) the three target domains in the stakeholder interviews, we next conducted a critical literature review. The purpose of the review was to identify and assess existing PROMs in the three domains, and narrow the list down based on predefined criteria (discussed below). We first conducted a search of PubMed and Google Scholar with search terms including, but not limited to, “patient reported outcomes rheum measures,” “patient reported experience of care measure,” “patient reported experience measure,” and “patient reported experience measure rheum.” We also searched relevant websites (eg, American College of Rheumatology) and reviewed the table of contents over several years for select journals (eg, Arthritis Care and Research). Articles were selected for review if abstracts described any type of PROM validity or reliability assessment. We examined each PROM and article for inclusion considering certain parameters (provided in Table 1), including whether validation was conducted with majority populations or arthritis-specific populations (33).

Our research study team, including leaders and staff in the Arthritis Foundation (EC, MSS, LM, MJ, AV, GE) and outside academic researchers (KES, RLB, EKZ), used results from the interviews and the critical literature review (more information in Results) to select three PROMs within each of the domains. Our criteria for review and selection included the strengths of the psychometrics, if the PROM needed input from medical providers that would not be possible to obtain through INSIGHTS, instrument burden (eg, how long it would take to complete), and whether there were fees associated with the PROM’s use (33). Lastly, the review considered whether the PROM items covered areas that the Arthritis Foundation could, from a practical standpoint, influence through targeted support, referrals, or programming.

Modified NGT. After the literature review, we developed and conducted three domain-specific, modified NGTs (34,35) to gather input and select a preferred PROM for each domain. We chose this method because NGTs are designed to reach consensus and allow input from each participant. We aimed to have 8 to 10 individuals per group based on NGT guidelines (34,36), including patients, caregivers, and one provider and one measurement expert per group to raise clinical and measurement considerations during NGT discussions.

Table 1. Literature review parameters examined and rated for each PROM

Methods, participants, and development of items
Participants: was the measure used with a general population, or has it been validated with patients with specific diseases, especially rheumatic diseases? How many participants were included in the stud(ies)?
Methods: how was the study carried out?
Development of items: what steps were included in the development of the items? The gold standard might include literature review; a panel, interviews, or focus groups with both experts and patients; revision of the measure; pilot testing (eg, content validity)
Administration
Frequency of Administration: how much time was in between administrations, both for test-retest purposes and for responsiveness to change purposes?
Time to Complete: did the article(s) include an estimate of how long it takes to complete the measure, either as a whole or by subscale? If so, how long?
Measurement
Measurement Properties: were key psychometrics evaluated (eg, reliability and validity), and if so, how should they be interpreted?
Items: how many items are included in the measure? If the measure has subscales, what are the subscales? What is the response scale?
Other relevant information
Registries in use: is the measure included on a well-known registry (ie, RISE, PR-COIN, VARA)? Was it previously reviewed in a literature review?
Associated Fees: what is the fee for using the measure and permissions required?

We recruited NGT patient or caregiver participants from a purposeful and snowball sample of the Foundation’s membership ($n \geq 100$), and we recruited provider and measurement experts based on their engagement with the Foundation on other initiatives. Potential participants were invited via email from Foundation staff (EC, MSS). We asked interested individuals to commit to both a prework survey and virtual NGT and used a survey to gather demographic information. Once participants were recruited, we used the demographic information to create three similar groups.

We modified the first step of the NGT process by substituting an online, prework survey using Qualtrics (37) in which participants reviewed and rated each of the three candidate PROMs in their group’s domain but were not told the names of the PROMs. Participants first rated the ease of answering each item and the importance of each item for individuals with arthritis. They then rated each PROM overall on its difficulty to complete, usefulness for tracking symptoms over time and comparing results to others, and degree to which results could assist the Foundation in improving its programs.

Within 2 weeks of the surveys, we conducted three virtual, domain-specific NGTs with one moderator (academic researcher) and one note-taker (Foundation staff member) at each session. Moderators (KES, RLB, EKZ) used domain-specific NGT facilitation guides and simple data displays of aggregated survey ratings to facilitate discussion related to each PROM, such as asking participants to explain their ratings and their thoughts on

Table 2. NGT and focus group participant demographics

	Response	NGT number (%)	Focus group number (%)
Gender	Female	17 (77%)	49 (89%)
	Male	5 (23%)	6 (11%)
	n	22	55
Age	Average age (and range)	57 years (38-75)	62 years (25-85)
	Average age of dependent (and range, n = 8)	-	15 years (10-19)
	n	20	56
Race (multiple answers allowed)	White	20 (95%)	35 (64%)
	Hispanic or Latino	-	1 (2%)
	Black or African American	1 (5%)	22 (40%)
	American Indian or Alaska Native	-	1 (2%)
	Asian	-	-
	Middle Eastern or North African	-	-
	Native Hawaiian or Pacific Islander	-	-
	n	21	55
Highest education	Less than high school	-	1 (2%)
	High school diploma/GED	-	5 (9%)
	Some college (including AA or technical degree)	3 (14%)	19 (36%)
	4-year college degree	7 (33%)	14 (26%)
	Graduate degree (eg, Masters, Doctorate)	11 (52%)	14 (26%)
	n	21	53
Role (multiple answers allowed)	Patient with a rheumatic disease	18 (86%)	48 (81%)
	Caregiver to a person with a rheumatic disease	1 (5%)	11 (19%)
	Measurement expert	2 (9%)	-
	Health care provider	1 (5%)	-
	Other	0	7 (12%)
	n	22	59
Diagnosis (self-reported or caregiver reported; checked all that applied)	Ankylosing spondylitis	2 (9%)	4 (7%)
	Fibromyalgia	2 (9%)	5 (9%)
	Gout	2 (9%)	9 (16%)
	Juvenile rheumatoid arthritis	0	7 (13%)
	Lupus	2 (9%)	6 (11%)
	Osteoarthritis	12 (55%)	31 (55%)
	Psoriatic arthritis	2 (9%)	2 (4%)
	Rheumatoid arthritis	6 (27%)	19 (34%)
	Other (eg, juvenile idiopathic arthritis, viral arthritis)	1 (5%)	10 (18%)
	n	22	56
Disease duration	<1 year	0	4 (7%)
	1-4 years	2 (9%)	19 (34%)
	5-10 years	4 (18%)	6 (11%)
	>10 years	13 (59%)	27 (48%)
	n	22	56

Note. A few participants did not respond to a few or all demographic questions.

the summary results. Following group deliberation, the NGTs culminated in asking participants to vote on the top PROM for each domain. Results of the NGTs were used to choose the final instrument per domain.

Focus groups. Following final PROM selection, we conducted six regional focus groups with additional patients and caregivers to obtain input on implementation of INSIGHTS, including messaging, recruitment, and sustaining participation. We

chose six focus groups based on research that this is sufficient to reach saturation (38) and on the desire to obtain input from individuals across different US regions. We established selection criteria to obtain diversity across diagnoses, disease severity, residence (eg, rural and urban), age, gender, and race/ethnicity. We recruited patients and caregivers through the Foundation's volunteer networks to reach a target of 8 to 10 participants per group. Three facilitators conducted the focus groups using a standard focus group guide and materials for participants, which included the selected PROMs and a brief overview of how the Foundation was planning to collect data. Prior to the focus groups, facilitators participated in a 1-hour training that included a review of the study background, the selected PROMs, and practice using the guide. Focus groups were recorded, transcribed, and analyzed by the academic researchers (KES, RLB, EKZ) using Dedoose and mixed inductive and deductive analysis (39,40).

Triangulation. Triangulation is a process used to cross-check data and assess saturation of findings. In this study, we used two of the four major types of triangulation described by Denzin (30): methodological triangulation and investigator triangulation. We used methodological triangulation by comparing results across different methods (interviews, NGTs, focus groups) and investigator triangulation by always having at least two academic researchers involved in data collection and analyses and involving other research team members in review and discussion of findings across our methods.

RESULTS

Overall, our mixed methods approach yielded input from 93 individuals, 69 (74%) of whom had some form of arthritis. Twelve individuals participated in the interviews; 22 participated in the survey portion of the NGT, of which 17 participated in the virtual NGTs (5-7 per group); and 59 (8-13 per group) participated in focus groups. Table 2 provides a description of our NGT and focus group participants; most (94%) were individuals with arthritis or caregivers of individuals with arthritis (eg, parents) who represented seven common types of arthritis. Although we set out to have one clinician and one measurement expert per NGT, we ended up with one clinician each in two NGTs and one measurement expert in the third NGT.

Stakeholder interviews. All stakeholders supported both the idea to create a system to collect PROMs over time (INSIGHTS) and the focus on the three domains identified in the Foundation's preliminary research. Participants identified multiple purposes for the initiative at different levels, including patients, the Foundation, and external stakeholders (eg, research community). Examples included personal use of results for tracking symptoms, and sharing with a medical provider; population-level reporting to identify areas in which to invest resources, service improvements, or advocacy by the Foundation; and data for research by external stakeholders. In relation to using data for research, less than half of the stakeholders

noted that attention would need to be paid to generalizability and accuracy of the data, such as verifying arthritis disease types.

When thinking about the development of INSIGHTS, top considerations reported by most participants included that data should be valuable and relevant for people with arthritis, data collection should be secure and confidential, and communications (eg, invitations) should provide a clear purpose for the initiative and pathway to participating. Participants' input on how to ensure value and relevance of INSIGHTS varied but included suggestions for how to message the program and provide individual data summaries that could be used by patients. Additional considerations reported by some participants included adding demographic questions, such as disease type and duration, and keeping track of events that may affect responses over time, such as surgery, medication, or lifestyle changes.

All participants supported the idea of longitudinal data collection, but no clear consensus emerged regarding frequency of data collection, although monthly or quarterly deployment appeared to have the most support. Some participants felt that the Experience of Care domain should be collected less frequently than the other two domains as health care visits happen less often.

Literature review. Our comprehensive literature review of potential PROMs revealed 37 existing domain-related PROMs with varying psychometric properties and number of items. These included 20 PROMs that assessed physical health, 11 that assessed social and emotional health, and 16 that assessed experience of care. A full list of the reviewed PROMs is provided in the article's supplemental material. Our research group review resulted in the selection of three PROMs that had subscales or short forms in both the physical health and social and emotional health domains: Psoriatic Arthritis Impact of Disease (PsAID-12) (14), 36-Item Short Form Survey (SF-36) (41), and the Patient-Reported Outcomes Measurement Information System (PROMIS)-29 with additional PROMIS Informational Support (version 2, four items) and Emotional Support short forms (version 2, four items) (42). We modified the PsAID slightly to be inclusive of all types of arthritis by replacing "psoriatic" with "your arthritis" in the instructions and items with this word. We felt these modifications would not significantly change the measurement properties for the PsAID, and we were committed to validate these modifications if the PsAID was chosen. Three additional PROMs were selected for the Experience of Care domain; Health Confidence questions (43), Health Care Empowerment Questionnaire (HCEQ) (44), and the Patient Self-Advocacy Scale (45). The selected PROMs were used in the modified NGT prework surveys and interviews.

Modified NGTs. Participants in the NGT groups ($n = 22$ survey, $n = 17$ virtual groups) provided input on the candidate PROMs and items. In general, the Physical Health domain group rated the PROMIS-29 and additional short forms as well as the PsAID higher in terms of ease of answering the items and importance.

After discussion, however, NGT participants settled on the PROMIS-29 and emotional support short form as being the best choice because of the format (ie, all answer options have values) and number of answer options (5 points versus 10 points). The Social and Emotional Health NGT group rated the PROMIS items and SF-36 similarly in terms of ease of answering the items and importance, but they clearly preferred the PROMIS items overall because of the ease of understanding, time frame (1 week versus 4 weeks), format of answer options, and wording. The Experience of Care NGT group had a more difficult time deciding on one PROM, rating the Patient Self-Advocacy Scale and Health Confidence questions higher than the HCEQ on ease and importance. During the virtual discussion, participants revealed that they liked the additional details asked in the HCEQ but had concerns about one of the subscales. By the end of the NGT discussion, participants decided that the HCEQ was better than the other instruments if the one concerning subscale was removed and questions were specifically connected to an arthritis-related health care visit. After reviewing the instrument psychometrics for the two remaining HCEQ subscales and follow-up discussions with a PROM measurement expert, we decided participants' suggested modifications to the HCEQ would not reduce its validity and reliability overall and that we could use the two subscales of the HCEQ.

During the NGT discussions, participants had additional general suggestions or concerns that were common across all three groups. These included suggesting that not all questions be asked each time to reduce burden, adding demographic questions to understand results, and mixed feelings about showing an individual's results compared with summary results from similar people. Participants suggested that some individuals may find it helpful, whereas others may feel worse if their results were not as good as others.

Focus groups. Participants (n = 59) reviewed the selected PROMs and offered perspectives on the implementation of INSIGHTS over time, including messaging, engagement and sustainment strategies, and frequency of data collection. Almost all participants expressed support for INSIGHTS, including the relevance of the domains and selected PROMs, and for longitudinal data collection.

Ideas for how to use the data were similar across all of the focus groups, including individual use to track personal trends, sharing with a provider to guide clinical decisions, informing Foundation programming, and advancing research. The top concerns participants raised about INSIGHTS included helping constituents see the "value add" for sustained engagement, ensuring data security and confidentiality, and making sure the purpose was clear. Some participants expressed concerns related to seeing personal results, particularly related to anxiety or depression, if compared with other people's results, whereas others felt that adding demographics, things about treatment (eg, medication), and other major life events (eg, divorce) were important to frame the results.

Similar to the stakeholder interviews and NGTs, there were mixed feelings about the frequency of administering INSIGHTS, although the most common recommendation was monthly. Focus group participants reiterated the idea that some things might change more frequently, such as physical function, or less frequently, such as experience of care, and should inform frequency.

Triangulation. As described in our methods, at least two members of the academic research team were involved in all analyses to achieve investigator triangulation (30). This included joint development of codebooks, review of coding, and identification of themes in the interviews and focus groups, as well as joint analysis of survey and qualitative data from the NGTs. Three members of our research team (EKZ, KES, GE) led the PROMs literature review and development of the criteria and ratings. All members of the research team reviewed results of the ratings and came to consensus on the final three PROMs in each domain for the NGTs.

We achieved methodological triangulation (30) by continually building on and comparing findings across methods (29), such as using findings from the NGT to develop the content of the beta INSIGHTS instrument for review in the focus groups. By comparing findings, we found substantial consistency in the main considerations and recommendations across our study steps. Table 3 summarizes the thematic findings per method that emerged in at least two and sometimes three methods. For instance, suggested uses for INSIGHTS data were very similar across all three methods. We also found that no clear consensus emerged regarding frequency of data collection, although monthly deployment appeared to have the most support across all three methods.

Overall. In the end, the final instruments for INSIGHTS included the PROMIS-29 Profile (v2.1), PROMIS Emotional Support Short Form (v2.0), and the HCEQ. Our study results also assisted the Foundation in deciding on the important demographic, disease-specific, and medication and lifestyle questions to include in INSIGHTS with minimal burden on participants. Initial ideas for what to include in this section came from the interviews, NGTs, and focus group results, which our academic research team used to draft initial questions. These were then reviewed and revised by the Foundation team (EC, LM, AV, GE) and clinician partners to finalize the items.

Our results also revealed the top considerations for initial implementation of INSIGHTS. The Foundation used this information in deciding to invest in a secure, multifaceted online platform for deployment of INSIGHTS, including customized delivery, messaging, and varied frequency per domain, and the potential to link responses over time and create individualized dashboards. At present, the decision to develop personalized data dashboards has been deferred pending analysis of early INSIGHTS data and review of options. However, the Foundation has created internal dashboards with de-identified aggregate data filterable by region/local volunteer "market," arthritis condition, gender, and age group

Table 3. Key triangulated themes across methods

Results	Interviews	NGTs	Focus groups
Usefulness of INSIGHTS			
Support for Foundation to implement INSIGHTS longitudinally and the three domains	X		X
INSIGHTS results can be used for multiple purposes by patients, Foundation, external stakeholders	X	X	X
Considerations for implementation			
Ensure and message value and relevance to patients	X	X	X
Vary frequency of data collection per domain to reduce burden	X	X	X
Communicate clear purpose and pathway for participating	X		X
Include demographics, medication, lifestyle changes to interpret findings and assist in future research	X	X	X
Choose items that are clear and applicable to ensure accuracy and quality of data	X	X	
Concerns			
Concerns about data privacy and security	X		X
Mixed feelings about comparing individual results to group results		X	X

to display data trends over time for use in programming decisions and potential future priorities for research. These dashboards use PROMIS T scores and cut points based on the summed raw scores for the items in each PROMIS measure or subdomain (eg, anxiety: four items, physical function: four items, etc) to establish symptom severity and level of functioning ranging from Within Normal Limits, to Mild, Moderate, or Severe symptoms (46). Gauge charts for each subdomain indicate the cumulative average score for all respondents, with color coding (green, yellow, orange, red) to indicate symptom severity or level of functioning based on T scores. Similarly, we used the summed raw HCEQ scores for each of the two subscales grouped into four categories of health care involvement, from “not at all” to “extremely” involved, but we hope to establish population norms in future work.

DISCUSSION

Our longitudinal mixed methods approach to ensure stakeholder involvement in developing INSIGHTS resulted in extensive patient input to ensure the program’s relevance for the arthritis community and appropriate use of PROMs outside of clinical settings through the Foundation’s patient networks. Our systematic approach led to the selection of the PROMIS generic PROMs over arthritis-specific measures. Although PROMIS has been shown to predict outcomes for, and has been validated with, individuals with arthritis (47,48), there are benefits and drawbacks to using generic versus disease-specific measures (49,50). When we presented both types of measures to NGT participants in the context of use for the INSIGHTS program, the benefits of PROMIS outweighed its lack of focus on arthritis. This may not be surprising given that we only offered one arthritis-specific PROM (PsAID) versus two generic PROMs (PROMIS and SF-36). However, given our goal to have valid PROMs that could be filled out without provider input (often required in many arthritis-specific PROMs) and the NGT consensus for PROMIS, the final choice of a generic PROM was supported.

Using the PROMIS tool also enabled us to leverage the T score-based cut points established by the HealthMeasures group to represent symptom severity and level of functioning by subdomain in internal dashboards, and potentially in future individual dashboards. Although these T scores are based on a normative, general population rather than our arthritis-specific population, they offered a starting point for characterizing trends in INSIGHTS data to guide programming decisions and other organizational priorities. We plan future analyses to validate the cut points for our arthritis population and contribute to the work already happening in this area (51) and to assess useability of this scoring for prioritizing programs and for individuals to monitor their own health status.

Our work also revealed a shortage of high-quality, experience-of-care measures that focus more on a patient’s initiative in getting care (eg, asking questions) versus actions taken by providers (eg, answering questions). The available measures do not appear to be widely used and were not validated for people with arthritis. Additionally, patients in our experience-of-care NGT had more trouble choosing the final PROM compared with patients in the other two NGTs. Thus, one of the goals for INSIGHTS is to collect and further test the chosen HCEQ items to assess their properties and usefulness for patients, the Foundation, and external stakeholders in the management of arthritis.

There are several limitations to our study. First, almost all of our participants had some association with the Foundation, which suggests some degree of interest or involvement in seeking information or trying to address difficulties with the disease, and potentially greater ability to access resources provided by the Foundation (eg, Internet access). However, because the primary purpose of INSIGHTS is to assist the Foundation in setting priorities and meeting the needs of their constituents, our purposive selection from Foundation membership is justified. We attempted to solicit input from a diverse group and achieved good success in some areas (eg, multiple diagnosis, 40% African Americans in focus groups). However, we lacked representation in other areas,

such as having a very low Hispanic and no Asian representation, fewer individuals with less-than-college education, and few males. This could lessen the relevance of INSIGHTS for certain groups.

Our literature review, although comprehensive, may have missed some potential PROMs, and our decision to exclude PROMs that required provider input limited the number of disease-specific options for the NGTs. However, the remaining PROMs had solid psychometric properties and were rated highly on other criteria (eg, reduced respondent burden), some of which emerged as important considerations for interview and NGT participants.

In the NGTs, we were not able to have a measurement expert and clinician in each group to provide their perspectives. Given the strong psychometric properties of the PROMs chosen to review in the NGTs and our purpose to create an instrument primarily for individuals with arthritis, we feel that our results would not have changed much with this additional input. Another potential limitation is the slight modifications of the HCEQ requested by our NGT participants. However, we plan to reassess the HCEQ properties once we have adequate data through the INSIGHTS program.

In the end, having an accessible, patient-driven measurement system backed by PROMs with strong psychometrics and selected with robust stakeholder engagement sets the stage for patients to more easily monitor their health both outside and inside health care settings and for patient networks, such as Live Yes!, to understand and respond to the needs of individuals living with arthritis. An additional potential strength of INSIGHTS, given the choice of the generic PROMIS measures, is that overall results and specific disease categories (eg, rheumatoid arthritis) can be compared with the general population over time.

INSIGHTS joins other initiatives, such as FORWARD (52), Arthritis Power (53), and PatientsLikeMe (54), to promote consumer ownership and engagement in improving health outcomes, disease management, and research for individuals with arthritis. Its distinction from these initiatives and existing registries is its broad reach through the Foundation's nationwide Live Yes! Network, which furthers the use and applicability of PROMs for on-the-ground programming, tailored support, and advocacy. INSIGHTS data are now being collected and used to support the Foundation's programming decisions and organizational priorities in research and advocacy. Future work will assess whether this innovative pairing of PROMs with a national, patient-driven network yields data-driven interventions, improved patient outcomes, and new research tools for other investigators.

ACKNOWLEDGMENTS

The authors would like to acknowledge the valuable participation and input of patients, providers, and other stakeholders as partners in this important work.

AUTHOR CONTRIBUTIONS

Drs. Schifferdecker and Knight and Ms. Butcher drafted and revised the article to include important intellectual content; Ms. Creek, Schrandt, Marrow, Jaffe, and Vinci and Dr. Eakin revised the manuscript to include important content; all authors approved the final version of the submitted article to be published.

Study conception and design. Schifferdecker, Butcher, Knight, Creek, Schrandt, Marrow, Eakin.

Acquisition of data. Schifferdecker, Butcher, Knight, Schrandt, Marrow, Jaffe, Vinci.

Analysis and interpretation of data. Schifferdecker, Butcher, Knight.

REFERENCES

1. U.S Food and Drug Administration. Patient-reported outcome measures: use in medical product development to support labeling claims. 2009. URL: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>.
2. Broderick JE, DeWitt EM, Rothrock N, Crane PK, Forrest CB. Advances in patient-reported outcomes: The NIH PROMIS. *EGEMS (Wash DC)* 2013;1:1015.
3. Weitzman ER, Wisk LE, Salimian PK, Magane KM, Dedeoglu F, Hersh AO, et al. Adding patient-reported outcomes to a multisite registry to quantify quality of life and experiences of disease and treatment for youth with juvenile idiopathic arthritis. *J Patient Rep Outcomes* 2018;2:1.
4. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2:1-74.
5. Beukelman T, Kimura Y, Ilowite NT, Mieszkalski K, Natter MD, Burrell G, et al. The new Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry: design, rationale, and characteristics of patients enrolled in the first 12 months. *Pediatr Rheumatol Online J* 2017;15:30.
6. Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Serv Res* 2013;13:211.
7. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract* 2006;12:559-68.
8. Santana MJ, Feeny D. Framework to assess the effects of using patient-reported outcome measures in chronic care management. *Qual Life Res* 2014;23:1505-13.
9. Dawson J, Doll H, Fitzpatrick R, Jenkinson C, Carr AJ. The routine use of patient reported outcome measures in healthcare settings. *BMJ* 2010;340:c186.
10. Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013;346:f167.
11. Barton JL, Katz P. The patient experience: patient-reported outcomes in rheumatology. *Rheum Dis Clin North Am* 2016;42: xv-xvi.
12. Wahl E, Yazdany J. Challenges and opportunities in using patient-reported outcomes in quality measurement in rheumatology. *Rheum Dis Clin North Am* 2016;42:363-75.
13. Roos EM, Lohmander LS. The Knee Injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes* 2003;1:64.
14. Gossec L, de Wit M, Kiltz U, Braun J, Kalyoncu U, Scivo R, et al. A patient-derived and patient-reported outcome measure for assessing psoriatic arthritis: elaboration and preliminary validation of

- the Psoriatic Arthritis Impact of Disease (PsAID) questionnaire, a 13-country EULAR initiative. *Ann Rheum Dis* 2014;73:1012–9.
15. Wolfe F, Michaud K, Pincus T. A composite disease activity scale for clinical practice, observational studies, and clinical trials: the Patient Activity Scale (PAS/PAS-II). *J Rheumatol* 2005;32:2410–5.
 16. Pincus T, Bergman MJ, Yazici Y, Hines P, Raghupathi K, Maclean R. An index of only patient-reported outcome measures, Routine Assessment of Patient Index Data 3 (RAPID3), in two abatacept clinical trials: similar results to Disease Activity Score (DAS28) and other RAPID indices that include physician-reported measures. *Rheumatology (Oxford)* 2008;47:345–9.
 17. Bartlett SJ, Orbai AM, Duncan T, DeLeon E, Ruffing V, Clegg-Smith K, et al. Reliability and validity of selected PROMIS measures in people with rheumatoid arthritis. *PLoS One* 2015;10:e0138543.
 18. Driban JB, Morgan N, Price LL, Cook KF, Wang C. Patient-Reported Outcomes Measurement Information System (PROMIS) instruments among individuals with symptomatic knee osteoarthritis: a cross-sectional study of floor/ceiling effects and construct validity. *BMC Musculoskelet Disord* 2015;16:253.
 19. Morgan EM, Mara CA, Huang B, Barnett K, Carle AC, Farrell JE, et al. Establishing clinical meaning and defining important differences for Patient-Reported Outcomes Measurement Information System (PROMIS®) measures in juvenile idiopathic arthritis using standard setting with patients, parents, and providers. *Qual Life Res* 2017;26:565–86.
 20. Pediatric Rheumatology Care and Outcomes Improvement Network. About PR-COIN. 2020. URL: <https://pr-coin.org/about>.
 21. American College of Rheumatology. RISE (Qualified Clinical Data Registry). URL: <https://www.rheumatology.org/1-Am-A/Rheumatologist/RISE-Registry>.
 22. Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* 2007;8:38.
 23. Kirwan JR, Bartlett SJ, Beaton DE, Boers B, Bosworth A, Brooks PM, et al. Updating the OMERACT filter: implications for patient-reported outcomes. *J Rheumatol* 2014;41:1011–5.
 24. Rose M, Bezjak A. Logistics of collecting patient-reported outcomes (PROs) in clinical practice: an overview and practical examples. *Qual Life Res* 2009;18:125–36.
 25. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf* 2014;23:508–18.
 26. Nelson EC, Dixon-Woods M, Batalden PB, Homa K, Van Citters AD, Morgan TS, et al. Patient focused registries can improve health, care, and science. *BMJ* 2016;354:i3319.
 27. Arthritis Foundation. Live Yes! Arthritis Network. URL: <https://www.arthritis.org/liveyes>.
 28. Callahan LF, Shreffler JH, Altpeter M, Schoster B, Hootman J, Houenou LO, et al. Evaluation of group and self-directed formats of the Arthritis Foundation's Walk With Ease program. *Arthritis Care Res (Hoboken)* 2011;63:1098–107.
 29. Schifferdecker KE, Reed VA. Using mixed methods research in medical education: basic guidelines for researchers. *Med Educ* 2009;43:637–44.
 30. Denzin N. *The research act: a theoretical introduction to sociological methods*. 2nd ed. New York (NY): McGraw-Hill; 1978.
 31. Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. *Field Methods* 2006;18:59–82.
 32. Dedoose, web application for managing, analyzing, and presenting qualitative and mixed method research data: version 8.2.14. Los Angeles, CA: SocioCultural Research Consultants, LLC; 2018.
 33. Knight E, Schifferdecker KE, Eakin GS. PRO measure profiling tool (PROMPT): the development of a rubric for optimal assessment of patient reported outcome measures in quality of life and additional domains. In: 26th Annual Conference of the International Society for Quality of Life Research; 2019 Oct 20-23; San Diego, CA; London, United Kingdom: Springer Nature; 2019. p. S19-20.
 34. Dobbie A, Rhodes M, Tysinger JW, Freeman J. Using a modified nominal group technique as a curriculum evaluation tool. *Fam Med* 2004;36:402–6.
 35. Rankin NM, McGregor D, Butow PN, White K, Phillips JL, Young JM, et al. Adapting the nominal group technique for priority setting of evidence-practice gaps in implementation science. *BMC Med Res Methodol* 2016;16:110.
 36. Varga-Atkins T, Bunyan N, Fewtrell R, Mclsaac J. *The Nominal Group Technique - a practical guide for facilitators*. Written for the ELESIG Small Grants Scheme. Liverpool, UK: University of Liverpool; 2011.
 37. Qualtrics software: August 2019. Provo (UT): Qualtrics; 2005.
 38. Guest G, Namey E, McKenna K. How Many Focus Groups Are Enough? Building an Evidence Base for Nonprobability Sample Sizes. *Field Methods* 2017;29:3–22.
 39. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int J Qual Methods* 2006;5:80–92.
 40. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005;15:1277–88.
 41. Ware JE Jr, Gandek B. The SF-36 health survey: development and use in mental health research and the IQOLA Project. *Int J Ment Health* 1994;23:49–73.
 42. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. Initial adult health item banks and first wave testing of the Patient-Reported Outcomes Measurement Information System (PROMIS) network: 2005–2008. *J Clin Epidemiol* 2010;63:1179–94.
 43. Wasson J, Coleman EA. Health confidence: an essential measure for patient engagement and better practice. *Fam Pract Manag* 2014;21:8–12.
 44. Gagnon M, Hibert R, Dubé M, Dubois MF. Development and validation of an instrument measuring individual empowerment in relation to personal health care: the Health Care Empowerment Questionnaire (HCEQ). *Am J Heal Promot* 2006;20:429–35.
 45. Brashers DE, Haas SM, Neidig JL. The Patient Self-Advocacy Scale: measuring patient involvement in health care decision-making interactions. *Health Commun* 1999;11:97–121.
 46. HealthMeasures. PROMIS® Score Cut Points. URL: <https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/promis-score-cut-points>.
 47. Witter JP. The promise of Patient-Reported Outcomes Measurement Information System—turning theory into reality. a uniform approach to patient-reported outcomes across rheumatic diseases. *Rheum Dis Clin North Am* 2016;42:377–94.
 48. Wohlfahrt A, Bingham III CO, Marder W, Phillips K, Bolster MB, Moreland LW, et al. Responsiveness of Patient-Reported Outcomes Measurement Information System measures in rheumatoid arthritis patients starting or switching a disease-modifying antirheumatic drug. *Arthritis Care Res (Hoboken)* 2019;71:521–9.
 49. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989; 27:S217–32.
 50. Wells GA, Russell AS, Haraoui B, Bissonnette R, Ware CF. Validity of quality of life measurement tools—from generic to disease-specific. *J Rheumatol Suppl* 2011;88:2–6.
 51. Nagaraja V, Mara C, Khanna PP, Namas R, Young A, Fox DA, et al. Establishing clinical severity for PROMIS® measures in adult patients with rheumatic diseases. *Qual Life Res* 2018;27:755–64.

52. Wolfe F, Michaud K. The National Data Bank for rheumatic diseases: a multi-registry rheumatic disease data bank. *Rheumatology (Oxford)* 2011;50:16–24.
53. Nowell WB, Curtis D, Thai M, Wiedmeyer C, Gavigan K, Venkatachalam S, et al. Digital interventions to build a patient registry for rheumatology research. *Rheum Dis Clin North Am* 2019;45:173–86.
54. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. *J Med Internet Res* 2010;12:e19.