## Research Report

# Patient Reported Outcome Measures in Adult Spinal Muscular Atrophy: A Scoping Review and Graphical Visualization of the Evidence

Jeremy Slayter[a,b], Lauren Casey[b] and Colleen O'Connell[a,b,*]
[a]*Dalhousie Medicine New Brunswick, Faculty of Medicine, Dalhousie University, Saint John, NB, Canada*
[b]*Stan Cassidy Centre for Rehabilitation, Horizon Health Network, Fredericton, NB, Canada*

**Abstract**.
**BACKGROUND:** Spinal Muscular Atrophy (SMA) is a hereditary neuromuscular disease with an estimated prevalence of 1/10 000 births. SMA is increasingly recognized as a multi-system disease with a need to study additional under-recognized health domains such as quality of life, fatigue, bulbar function, respiratory function, and independence.
**OBJECTIVE:** Identify and assess reported evidence from the literature investigating Patient Reported Outcome Measures (PROMs) in adults with SMA. Develop a novel method drawing from network theory to graphically depict the literature, PROMs, and supporting psychometric evidence.
**METHODS:** A scoping review was completed following PRISM-ScR, COSMIN and JBI scoping review guidelines. Literature investigating PROMs in adult SMA or neuromuscular disease was identified from peer-reviewed and grey databases. A network graph was derived from extracted data.
**RESULTS:** 5292 articles were retrieved, 81 articles met inclusion criteria; corresponding to 31 unique PROMs. Only two PROMs were developed specifically for SMA. Few PROMs covered multiple domains of health. Most PROMs were incompletely validated, focusing on concurrent validity, and few assessed responsiveness or internal consistency.
**CONCLUSIONS:** PROMs are emerging tools for monitoring and assessing adults with SMA. Despite their potential benefits, additional validation studies should be completed prior to their use for clinical decision-making. Network graphics may represent a technique to aid in the visualization of evidence supporting a scoping review.

Keywords: Review, muscular atrophy, spinal, patient reported outcome measures, patient outcome assessment, rehabilitation, neuromuscular disease, reproducibility of results, health status, psychological tests, neurodegenerative diseases

## INTRODUCTION

Understanding the many impacts of disease is essential to providing optimal patient care [1, 2]. One of the many tools available for clinicians and researchers to objectively quantify patient experiences are Patient Reported Outcome Measures (PROM). PROMs are defined by the Canadian Institute for Health Information (CIHI) as measurement instruments which are completed by patients, and which obtain information on aspects of the patients' health status relevant to domains such as quality of life, symptoms, function, pain, and physical, mental or social health [3]. PROMs are known to be valuable tools capable of capturing patient changes which may otherwise be missed [1, 2, 4–6]. PROMs are rapidly

*Correspondence to: Colleen O'Connell, Stan Cassidy Centre for Rehabilitation, 800 Priestman Street, Fredericton, New Brunswick E3B 0C7, Canada. Tel.: +1 506 447 4294; Fax: +1 506 447 4749; E-mail: drcolleenoconnell@horizonnb.ca.

being studied and implemented in areas of medicine including oncology, orthopedics, mental health, and chronic disease management [1, 2, 7–9]. PROMs are also being required by regulatory agencies [10]. PROMs prompt self-reflection from the patient and provide patients with an opportunity to raise concerns with clinicians, and may improve clinician awareness of the breadth of patient challenges [4]. Despite their utility, a major challenge to the uptake of PROMs is limited clinician exposure [2]. Secondly, there may be limited existing literature investigating PROMs in the population of interest [6]. PROMs used in other populations have often been incompletely studied or have limited literature to support their use [4–6, 11, 12]. Understanding the status of the body of literature of PROMs within the population of interest is important when considering the inclusion of PROMs into standard clinical monitoring and assessment practices. For a PROM to be useful, it must be valid, responsive, reliable, and reflective of the patient experience [11, 13–16]. Incorporating PROMs with strong supporting evidence into routine clinical practice can improve the understanding of how medical treatments impact patients, and can inform future designs of clinical trials with more clinically meaningful data capture [1, 5].

Spinal Muscular Atrophy (SMA) is a hereditary neuromuscular disease with an estimated prevalence of 1/10 000 to 1/11 000 births [17, 18]. SMA is characterized by the progressive loss of lower motor neurons, leading to several symptoms including progressive flaccid paralysis [19]. There are at least 33 causative genes identified [20]. 95% of SMA is due to a mutation in the Survival of Motor Neuron 1 (SMN1) gene [17]. A homologue of SMN1, SMN2, modulates the clinical severity of SMA by producing a low-functioning form of the SMN protein [20]. The copy number of SMN2 generally correlates to the types of SMA described clinically [20, 21].

In 2018, a landmark randomized clinical trial, Mercuri et al. 2018, was published, beginning a shift in the therapeutic landscape of SMA, which has now expanded to multiple therapeutic options for individuals with SMA [22, 23]. Despite these therapeutic advances, most clinical trials have been conducted in a pediatric SMA population. It is continued to be debated whether adults with SMA are receiving significant benefits from these treatments which often come with a high financial cost [24]. A key component to advancing this discussion is the ability of PROMs to capture alternative but valid and mean-ingful outcomes which more closely align with the patient experience [23, 24].

Patients, clinicians, and researchers are increasingly recognizing SMA as a multi-system disease and the need to study neglected outcomes such as fatigue, bulbar function, communication, respiratory function, and functional independence [13, 23–29]. The need for additional outcomes has arisen as a result of patient reported changes in function which were not captured by standard monitoring [5, 13, 23, 24, 26–30]. Despite this increased recognition, most studies have continued to rely on traditional measures, and few have incorporated PROMs into research designs [31]. Two previous reviews included PROMs available for use in SMA, although neither review focused on adults or PROMs in particular, and neither included validation data [13, 32].

A scoping review and network visualization was conducted with the purpose of investigating, characterizing, and visualizing currently existing PROMs which have been used in adult SMA and similar neuromuscular diseases (NMD). In addition, the review sought to describe the current trends of use and domains of function being evaluated to highlight areas where further research is needed. This review was not intended to evaluate or select an optimal PROM for use in adult SMA.

## MATERIALS AND METHODS

### Review & Search Methodology

This review utilized a scoping methodology to assess available literature examining PROMs in adult SMA. To improve the transparency, reporting, and applicability of the review; the Joanna Briggs Institute (JBI) scoping review, PRISMA-ScR, and COSMIN guidelines were followed [14, 33, 34]. The PRISMA-ScR and JBI guidelines both aim to standardize the reporting and methods of scoping reviews [34, 35] while the COSMIN guidelines pertain specifically to systematic reviews of PROMs; however, it was felt that many of the COSMIN guidelines were valid for the purposes of this review [14]. There is no previously published protocol for this scoping review.

For inclusion, PROMs had to meet the definition set by the CIHI, and had to be used in an adult SMA or similar NMD population [3]. Given the similarity in disease profile and in monitoring and management, a study which used an NMD population that included SMA was also determined to be appropriate to include in this review. In the case where minimal

validation data was available for a PROM, an alternative population could be included, provided all other inclusion criteria were met. As a result, the inclusion criteria required that articles must have evaluated PROMs which were:

1. Studied in an adult SMA, NMD population OR be in an alternative population in the instance where there is limited validation data available;
2. Completed by the patient (and/or option for caregiver to complete);
3. Informed on aspects of the patients' health status;
4. Relevant to domains such as quality of life, disease symptoms, function, pain, physical health, mental health or social health.

Included PROMs must have been used in an article which was published during or after 2016. The date of publication requirement was determined to represent the modern therapeutic, monitoring and assessment practices of researchers and clinicians. Prior to 2016, most studies focused on physical function and outcome measure sensitivity to change, and alternative non-motor endpoints were not strongly considered until after therapeutic options were being introduced for clinical use [20]. Original PROM sources were also included in the review when available. PROM characteristics included in the review were forms of validity, reliability, responsiveness, Rasch analysis, and others (e.g., time to completion, number of questions, previous use in clinical trials, proprietary status).

If inclusion criteria were met, any type of study including experimental, observational or review was considered. Grey literature could also be included if inclusion criteria were met. Articles unavailable in English were excluded. Support from institutional libraries (Horizon Health Network, Dalhousie University, and University of New Brunswick) and/or contacting article authors was used to locate any remaining full-text articles, published data, or if additional data was required. If full text articles or necessary data continued to be unable to be obtained, the article was then excluded.

The search strategy utilized PubMed, Google Scholar, Scopus, Embase, Web of Science, and grey literature to identify articles. Medical Subject Heading (MeSH) keywords were used when available. For non-MeSH databases, permutations of both MeSH and synonym keywords were used to ensure optimal database querying (Fig. 1). In the instance where a search produced an excessively large number of

**Search Terms:** *Spinal Muscular Atrophy, Muscular Dystrophy, Neuromuscular Disease, Motor Neuron Disease, Adult, Adolescent, Patient Reported Outcome Measures, Patient Outcome Assessment, Surveys and Questionnaires, Activities of Daily Living, Disability Evaluation, Meaningful.*

Fig. 1. Search Terms.

results due to database search algorithms, results were reviewed until there were consistently no longer any relevant articles identified.

The literature search was completed in June 2021. Search results were initially screened (JS) before being imported into Zotero and cataloged by database of origin if inclusion criteria were felt to be met [36]. After initial screening, all abstracts were then reviewed and further categorized based on PROMs used. Full text articles were then reviewed for inclusion criteria, and data was extracted to a comprehensive PROMs Table (Supplementary Table 1). During the full text review, forward and backward reference searching was used to identify any additional articles which were not captured during the initial search.

*Data Extraction*

This review extracted various psychometric properties of included PROMs as recommended by COSMIN [14]. COSMIN further recommended the evaluation of psychometric properties; however, the intention of this review was not to evaluate or determine an optimal PROM, but instead to compile the available evidence to inform clinicians and researchers. Hence, the COSMIN risk of bias checklist, Quality Criteria, and GRADE assessments were not applied as part of this study [14].

Measures of internal consistency, reliability, measurement error, responsiveness, and Rasch analysis were extracted. *Internal Consistency* (IC) describes how well several items in a scale vary together in a sample [37]. IC is often measured with Cronbach's Alpha, or Pearson's Correlation Coefficient (PCC) [14]. *Reliability* refers to the repeatability of a measure over various conditions [37, 38]. Three types of reliability were extracted for this review: test-retest reliability (TRT), inter-rater reliability, and intra-rater reliability. Reliability is typically reported with PCC, or Intraclass Correlation Coefficient (ICC). TRT reliability measures the stability of responses over time with the same respondents. Inter-rater reliability

assesses the stability between two raters at one point in time, while intra-rater reliability assesses the stability between instances of rating by one rater at one point in time. *Concurrent Validity (CV)* is a measure of how well the scale measures the underlying latent construct by comparison to a gold standard test [37]. CV is measured with PCC or Spearman's Ranked Correlation (SCC). *Responsiveness* is a measure of the sensitivity for a scale to detect changes [39]. The most common methods to measure responsiveness include the T-statistic, Effect Size (ES), Guyatt's Responsiveness Statistic, and standardized response mean (SRM) [40]. Other forms of responsiveness are in relation to clinically important changes (i.e., Minimal clinically important difference [MCID]), or expected variance due to measurement error (i.e., standard error of measurement [SEM], and minimally detectable change [MDC]). Lastly, *Rasch Analysis* is a scale assessment technique that examines the extent that observed data are in accordance to a Rasch model. Rasch Analysis seeks to optimize the validity and simplicity of a measure [41–43]. Combined, these psychometric properties can be used to describe how well a particular PROM reflects an underlying construct, if the PROM is able to detect changes, and how much error occurs during its use [14, 44].

Additionally, other basic PROM characteristics were extracted when available (e.g., number of Items, time to completion, option for caregiver completion, proprietary status). Six domains of function were created by grouping together themes previously identified as important in persons with SMA by Mongiovi et al. and Osmanovic et al. [28, 45]. The themes were grouped into the following domains:

1. *Physical Function:* Assessing aspects of physical function and mobility;
2. *Mental Health & Cognition:* Assessing aspects of emotional or mental health, social roles, and cognition;
3. *Fatigue*: Assessing aspects of Fatigue, tiredness or sleep status;
4. *Communication & Speech*: Assessing aspects of Bulbar function, speech, communication, and sight;
5. *Pain*: Assessing pain of any type;
6. *Systemic Issues*: Assessing any other signs of systemic illness, or which provided a chance to capture aspects which were otherwise not covered.

For a PROM to be marked as assessing a domain, the PROM must have addressed at least one aspect of function held within that domain as defined above.

*Graphical Network Visualization of PROMs*

A novel method of visualizing a literature review was built by creating an adjacency matrix from derived from the comprehensive PROMs Table (Supplementary Table 1), prior to being importing into R for creation with *igraph*, *ggraph*, and *ggplot2* packages [46–50]. The layout algorithm *stress* was selected as it best visualized the relationships between PROMs, characteristics, and articles. The *stress* algorithm is a widely used technique which seeks to minimize the stress function [51]. Centrality, a measure of relative importance, was calculated using the Katz centrality method [52, 53]. The size of nodes represents centrality, while node shapes represent each of the patient populations when applicable. Lastly, node colour represents the type of node (e.g., article vs PROM vs characteristic). The network graphic examines the relationships between supporting evidence of measures, how evidence evaluated PROMs, and how PROMs evaluate patients based upon domain.

The threshold to determine a relationship existed for the network graph was any mention of the topic directly (e.g., physical function domain must have included some question pertaining to physical function, while concurrent validity must have been evaluated by some means within the study). This threshold was chosen to recognize that there is no well-defined standard of assessment for most measure characteristics. Therefore, the network graphic allows for the recognition that each of these studies have different methodologies and evaluated characteristics and domains with variable methods.

*Statistical Analysis*

Through the compiled PROMs table, descriptive statistics (e.g., sum, mean, minimum, maximum) were obtained to describe the status of the literature, frequency of PROM use, and frequency of psychometric evaluation. Additional forms of descriptive statistics were obtained through the adjacency matrix to develop the network graph (e.g., centrality). No hypothesis testing was completed during this review.
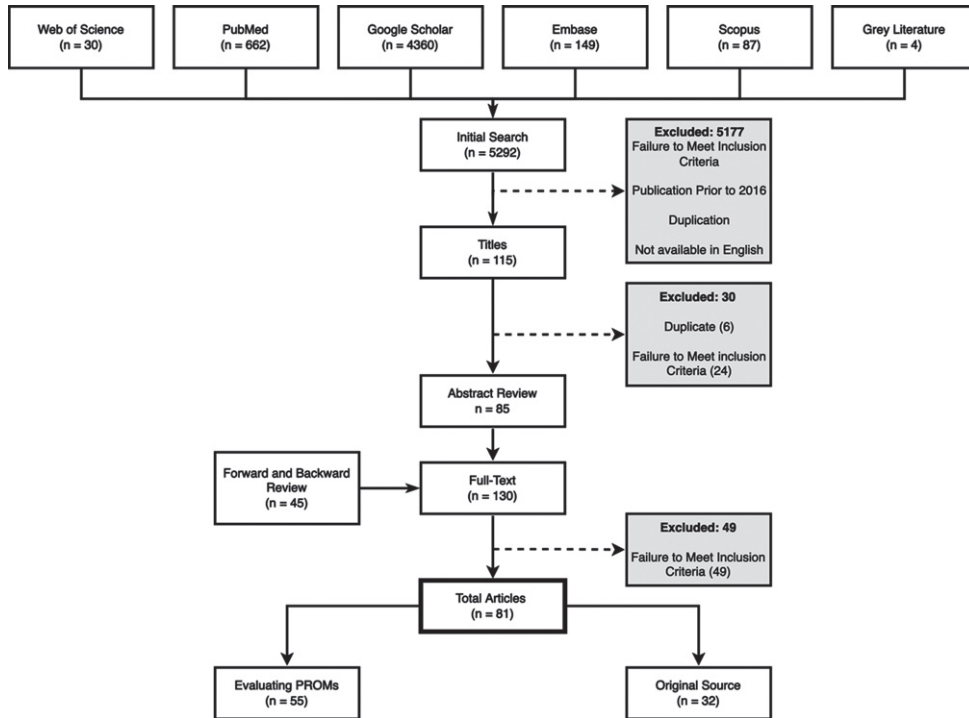
Fig. 2. Flowchart of review methods.

## RESULTS

### Search Results

There were 5292 total articles identified through the review process (Fig. 2). 5288 were found through scholarly databases, and an additional 4 grey literature articles were identified through non-academic databases. 115 articles met inclusion criteria prior to abstract review. After abstract review, 30 articles were excluded. During the full text review, forward and backward searching revealed an additional 45 articles meeting criteria. After full text review of all remaining articles, a further 49 articles were excluded after determining that they failed to meet inclusion criteria. In total, 81 articles were included in the final review and subsequent synthesis.

Of the 81 articles included in the final review, 77 were peer-reviewed articles and 4 were grey literature (Fig. 3, Table 1). 32 articles were original sources that were the first instance of publication of a PROM, while 55 articles directly evaluated PROMs. Of the 4 grey literature articles, each was an original PROM source from the general population. 46 articles evaluated psychometric data, with some articles assessing more than one aspect of psychometrics. Most studies

Table 1
Population Breakdown by Article Type

| Article Type | Population | | | |
|---|---|---|---|---|
| | Total (n) | SMA | NMD | Rehabiliation | General |
| Total | 81 | 21 | 21 | 6 | 33 |
| Peer-Reviewwd Articles | 77 | 21 | 21 | 6 | 29 |
| Grey Literatue | 4 | 0 | 0 | 0 | 4 |
| Original Source | 32 | 2 | 3 | 3 | 24 |
| Psychometric Evaluation | 46 | 14 | 18 | 5 | 9 |

evaluated concurrent validity (37) and internal consistency (22). Fewer evaluated test-retest reliability (17) or responsiveness (9) (Fig. 4).

### PROM Results

A total of 26 PROMs were initially identified. 5 were subsequently excluded for failing to meet inclusion criteria. An additional 10 were identified through forward and backward reference searching. 31 PROMs are included in the final review and synthesis (Supplementary Table 1).

Most PROMs evaluated domains of Physical Function (25, 80%), Mental Health and Cognition, (21,
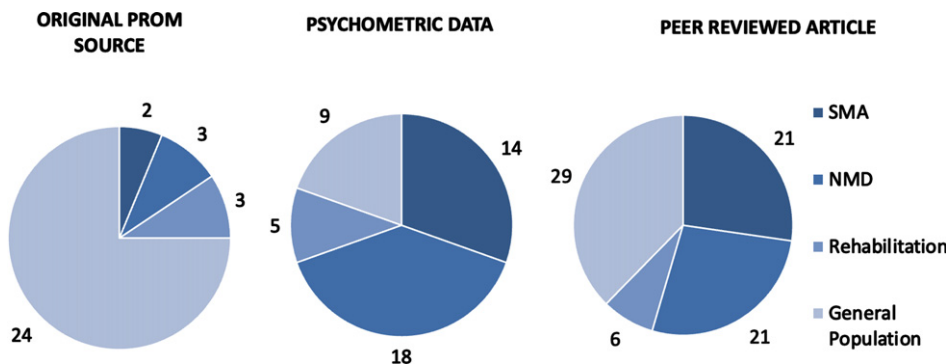
Fig. 3. Population breakdown.

67%), and Fatigue (17, 54%). Fewer evaluated Communication and Speech (13, 42%), Pain (11, 35%) or Systemic Issues (9, 29%) (Fig. 4). Only 7 PROMs could evaluate all domains of function. All included PROMs were administered through an interview with the patient. 3 PROMs had a caregiver (or proxy) administration option validated. 87% of PROMs were available without cost. The Activity Limitation Measure (ACTIVILIM), Quality of Life – Neuromuscular Disorders (QoL-NMD), and the Egen Klassifikation 2 (EK2) each had a completed Rasch analysis in either an SMA or NMD population. Only the SMA Independence Scale (SMAIS-ULM) and SMA-Health Index (SMA-HI) were developed specifically for SMA, all remaining PROMs were originally developed for a broader population and were re-purposed to fit an SMA population.

The most frequently cited PROMs were the Short Form 36 (SF-36) [14], Patient Reported Outcome Measurement Information System (PROMIS) [11],

and Fatigue Severity Scale (FSS) [10] (Fig. 5). Across all PROMs, there was an average of 5 articles per PROM. The maximum number of studies evaluating a particular PROM was 14 (SF-36), and the least was 1 (Self-care in Motor Neuron Disease Index [SCNMD]). Most studies evaluated 2 PROMs (mean of 1.90) at once. While 3 reviews were identified investigating outcome measures in an NMD or SMA population, none focused on adult SMA or PROMs in SMA.

*Network results*

The network graph and visualization of the evidence (Fig. 6) demonstrates the relationships between articles, PROMs, and characteristics visually. The network consists of 123 nodes which include each article, PROM, characteristic, and domain of function. The nodes are connected by 379 edges representing each relationship connection nodes.
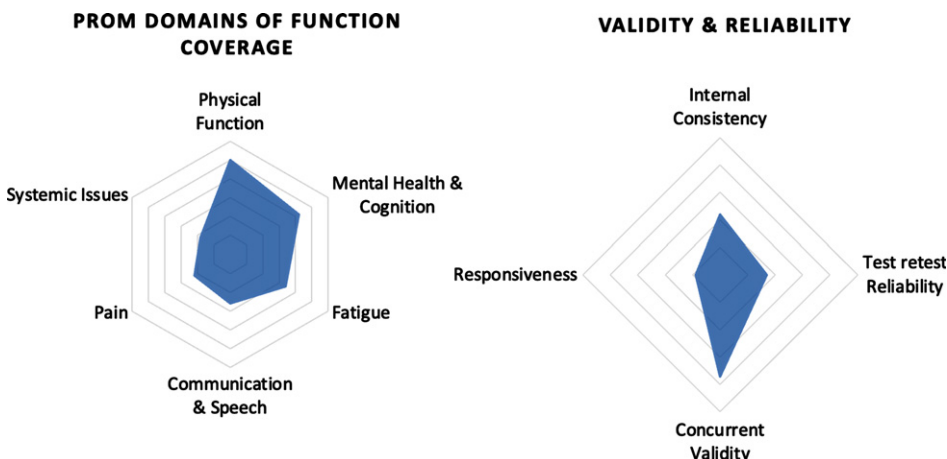


Fig. 4. Spider chart of psychometrics & domain of function.
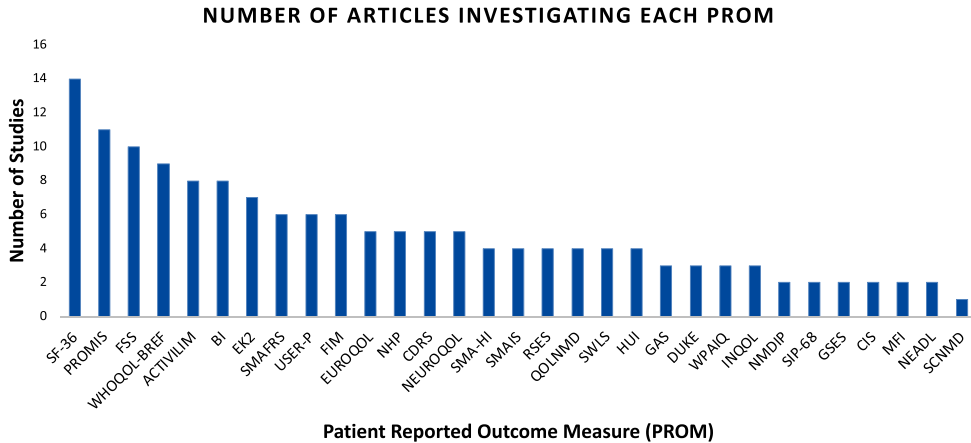
**NUMBER OF ARTICLES INVESTIGATING EACH PROM**



Fig. 5. Article Count Bar Chart.
SF-36 - Short-Form 36; PROMIS – Patient Reported Outcomes Measurement Information System; FSS – Fatigue Severity Score; ACTIVILIM – Activity Limitation Measure; BI – Barthel Index; EK2 – Egen Klassification 2; SMAFRS – Spinal Muscular Atrophy Functional Rating Scale; USER-P – Utrecht Scale for Evaluation of Rehabilitation Participation; FIM – Functional Independence Measure; EUROQOL – EUROQOL 5D/5L Measure; NHP – Nottingham Health Profile; CDRS – Connor Davidson Resiliency Scale 10; SMA-HI – Spinal Muscular Atrophy Health Index; SMAIS – Spinal Muscular Atrophy Independence Scale; RSES – Rosenberg Self-Esteem Scale; QOLNMD – Quality of Life Neuromuscular Disorders; SWLS – Satisfaction with Life Scale; HUI – Health Utility Index 3; GAS – Goal Attainment Scale; DUKE – Duke Health Profile; WPAIQ – Work Productivity and Activity Impairment Questionnaire; INQOL – Individualized Neuromuscular Quality of Life Questionnaire; NMDIP – Neuromuscular Disease Impact Profile; SIP-68 – Sickness Impact Profile 68; GSES – General Self Efficacy Scale; CIS – Checklist of Individual Strength; MFI – Multidimensional Fatigue Inventory; NEADL – Nottingham Extended Activities of Daily Living Scale; SCNMD – Self-Care in Motor Neuron Disease Index.

There are two central groupings of nodes, the first is the measure characteristic grouping, and the second is the domain of function grouping. Both groups are a function of two levels of data being represented within the graph. The measure characteristic group holds the article level data and their accompanying psychometric characteristics. The domain of function group holds the PROM level data, which also includes psychometrics and domains of function. The PROMs themselves are found to be interspersed throughout the network based upon their connectivity to both groups.

## DISCUSSION

### Interpreting the network graphic

The network graphic (Fig. 6) provides a visual representation of the status of the literature examining PROMs in adult SMA. The network graphic shows that there are a smaller number of PROMs which are most connected to other nodes; these PROMs (SF-36, USER-P, FSS, FIM, EK2, ACTIVLIM, SMAFRS) are situated in between both groups. The PROMs which are in the periphery, are new and upcoming, are rarely used, or have relatively little evidence to support their use. The peripheral PROMs may represent an opportunity for further research studies, as there is likely much that the literature does not know about their use in an adult SMA population. By utilizing both the network graphic (Fig. 6), and the comprehensive PROMs table (Supplementary Table 1), one can visually assess the level of support of PROMs and status of the literature and review the extracted objective evidence from each article.

### The current role of PROMs for adults with SMA

PROMs for adults with SMA is a rapidly evolving area due to the therapeutic advances and the recognized importance of PROMs in capturing patient experiences. With a total of 31 identified PROMs, the number of available PROMs is quickly growing. Most currently existing PROMs target physical function (80%), while few PROMs can capture aspects of every domain. This result is expected as the focus of research in SMA has traditionally been focused on physical function, with only recent recognition of additional impacts of the disease. Few PROMs evaluate domains of sexual function, speech, and other systemic issues, despite evidence supporting their importance [28, 45]. New PROMs should seek to further incorporate poorly captured domains. In pursuit
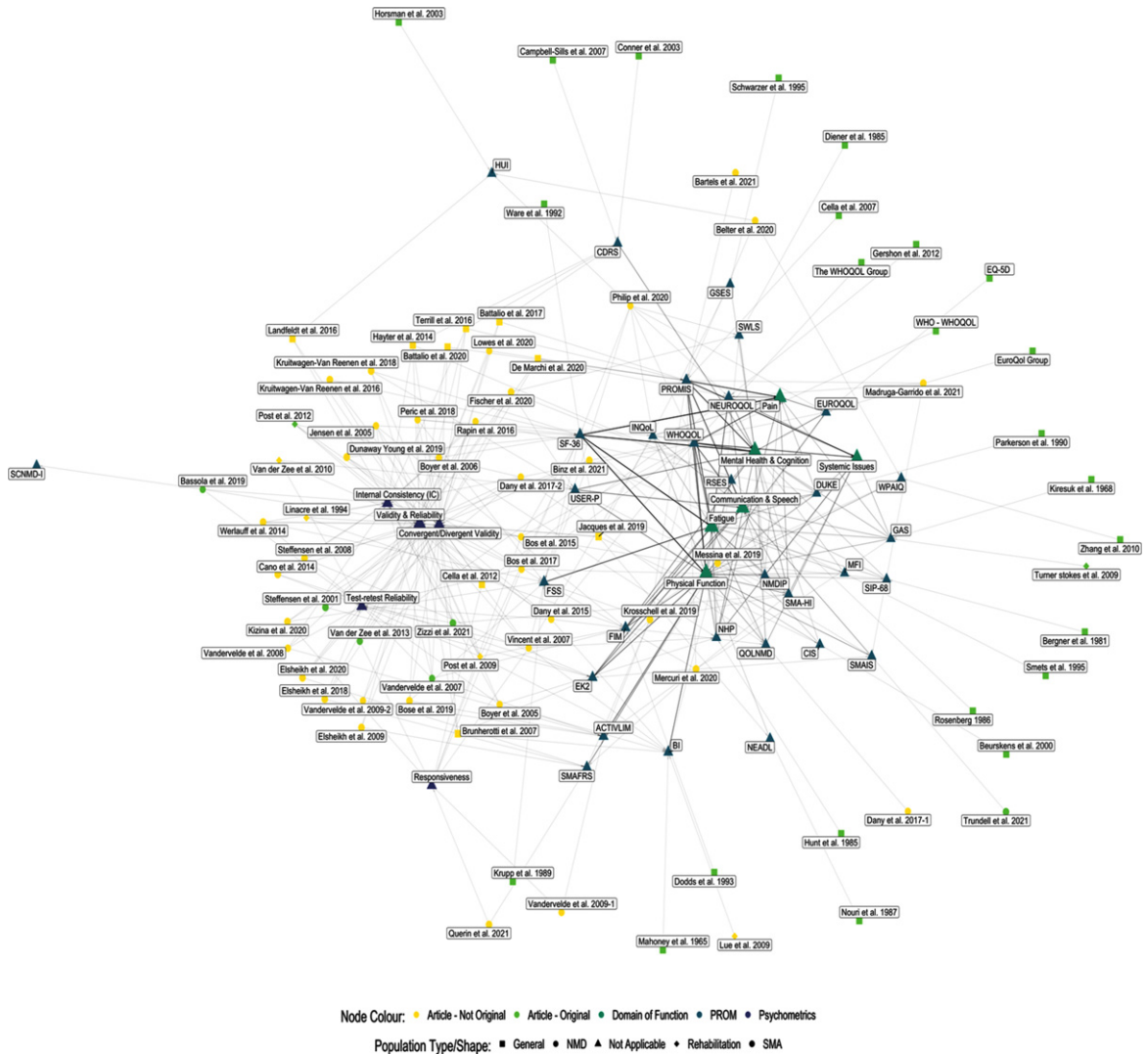
Fig. 6. Network graph.

of developing an optimal PROM in adult SMA, it is important to maintain open communication with all healthcare stakeholders, and to continue to critically assess emerging psychometric and clinical evidence so that PROMs may continue to adapt into the "fit-for-purpose" role required of them by patients, clinicians, and researchers [5].

Many of the identified PROMs are not well established or validated, with the most used PROMs continuing to have gaps in psychometric evidence. Most PROMs have been repurposed to fit an SMA population due to the traditional monitoring environment not requiring measures to be particularly responsive, and instead were used to

broadly categorize function. PROMs like the SF-36, WHOQOL-BREF, and the Functional Independence Measure (FIM) are frequently used and well supported by psychometric evidence, but it is possible that their content validity is sub-optimal in adult SMA due to their general intended use. These measures may be able to discriminate between adults with SMA, but it is largely unknown if they will be responsive to the subtle changes in function which have become particularly important in the modern therapeutic era [24]. Most PROMs evaluate a portion of the domains of assessment, but only 7 *could* evaluate all domains. Included in this group, are the PROMIS and NEUROQOL scales which are modular PROMs that

have many subscales available for use. Of the PROMs capable of measuring all domains of assessment, only the SMA-HI and the NMDIP were specific to NMD, with the remaining 5 PROMs being applied to SMA and/or NMD but developed for a more general population.

There is a tendency for articles to focus on concurrent validity (Fig. 4), rather than other equally important measures of validity, reliability, or responsiveness. This focus on concurrent validity may lead to an over-estimation of the true level of validation that a PROM has achieved in the adult SMA population. Additional study of other forms of validity, responsiveness, and reliability will provide more in-depth understanding of how PROMs capture patient experiences.

One of the most used PROMs is the SMAFRS which was initially developed from the Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS) and WeeFIM protocol to facilitate discrimination of SMA disease severity based upon SMN2 copy number [54]. Despite the original intentions of the SMAFRS, it has quickly become a standard tool in in assessing function, appearing in multiple recommendations, and clinical studies [25, 55–57]. Despite its widespread use, the SMAFRS continues to have limited validation and psychometric data supporting its use, and unlike its relative the ALSFRS, does not include assessment of speaking or respiratory function.

Having strong psychometric evidence is important to establish how a PROM will react in a research setting. Without this evidence, there is an increased risk of committing Type I and/or Type II errors when using PROMs which are not comprehensively validated [58]. To remedy this risk in an adult SMA population, thorough investigation of PROMs should be undertaken prior to their implementation in clinical decision-making [6].

*What's next for PROMs?*

Newly developed PROMs specific to SMA may provide a solution to the current patchwork quilt of validation and utilization. New upcoming measures such as the SMA-HI, ACTIVLIM, and the SMAIS may represent the next generation of PROMs in adult SMA, despite a current lack of validation data [59–61]. All three were developed for SMA or NMD to monitor disease progression in the new therapeutic environments and were intention-ally developed to optimize validity, reliability, and responsiveness. Since the completion of this review, the SMAIS and NEUROQOL have had further validation studies completed in an SMA population [61, 62].

Further validation efforts may become more challenging owing to the changing therapeutic environment with increasing numbers of persons with SMA receiving disease-modifying treatments which will require careful study design and patient selection. The careful design, development, and validation of PROMs in SMA, in collaboration with persons with lived experience, will maximize the likelihood of producing a sensitive, valid, and repeatable measure which is reflective of the patient's experience. If these upcoming measures can continue to show promising validation data, be easily interpretable and implementable for clinicians, and can be incorporated into research studies, they may be an important part of the future in assessing and monitoring adults with SMA.

Another key factor for encouraging the uptake of new, robust PROMs is to be accessible by both the research and clinical communities. If PROMs are unable to be located, understood, or affordable; they will not be widely adopted.

PROMs represent an opportunity to provide valid, responsive, and repeatable data that is both comprehensive and reflective of the experiences of persons with SMA [5, 11, 13–16]. With further investigation, currently existing PROMs may be found to be appropriate for use or it may be possible that stitching together the best properties of currently existing PROMs may lead to new measures which better capture a person's ability, while still being strongly supported by psychometric evidence. Until a robust PROM becomes available, it is important to be cautious when relying upon PROMs to draw significant conclusions until it is understood how PROMs capture the information which is most important to patients and clinicians.

*Limitations*

This review falls between a scoping review and a systematic review as the body of literature investigating PROMs in adult SMA populations continues to be relatively small. It is likely that this review included most of the available relevant literature. In the future, as the amount of literature increases, a more robust systematic review and meta-analysis may be war-

ranted which adheres strictly to COSMIN guidelines and seeks to determine which PROM is best supported. If a systematic review was completed today, there would be a risk of introducing a 'search satisficing' bias due to the finding that most PROMs do not have a complete validation available [63]. This bias could inadvertently lead to inappropriate PROM selection, particularly when the PROMs with most evidence are known to be imperfect. When a sufficient body of evidence exists, an international consensus guideline, systematic review, or an alternative approach to harmonize and unify PROMs used for adults with SMA would be highly desirable and clinically valuable. Another limitation of this study is that there was only one person (JS) who conducted the initial review. Optimally there would be two or more reviewers; However, due to there being only one reviewer, it is likely that the review criteria were more consistently applied across all articles because of only one interpretation of the criteria.

The network graphic and characteristics are both optimistic estimates of the relationships and evidence of PROMs. It is likely that PROMs are less robustly supported, or incompletely cover domains of function than described by the network graphic. It is argued that despite this, it is reasonable as this review did not seek to evaluate the quality of articles identified or comprehensiveness of domains of function. Furthermore, there is no saturation of data, and despite the optimistic estimate many gaps in evidence remained. It is possible that some PROMs do not adequately assess the domains of function, or that the methodology used to validate the PROM was inappropriate for reasons of introduced bias, poor study design, or others. Future research should seek to evaluate the comprehensiveness of PROMs to cover various domains of function.

*Conclusion*

This review sought to outline the status of the literature on PROMs in adult SMA. This review can be used to support the development and refinement of further research, and the clinical implementation of PROMs. A variety of PROMs already exist, but few are robustly researched or comprehensive in scope. New PROMs are being developed but most have not yet come to maturity and have not yet been widely adopted or translated to practice. As new PROMs become available, it is important to continue to evaluate and validate the evidence to understand both their theoretical and practical value in monitoring and assessing adults with SMA.

## SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/JND-221595

## REFERENCES

[1]    Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. BMJ. 2015;350:g7818. https://doi.org/10.1136/bmj.g7818

[2]    Black N. Patient reported outcome measures could help transform healthcare. BMJ. 2013;346:f167. https://doi.org/10.1136/bmj.f167

[3]    Patient-reported outcome measures (PROMs) | CIHI n.d. https://www.cihi.ca/en/patient-reported-outcome-measures-proms (accessed June 13, 2021).

[4]    Greenhalgh J, Gooding K, Gibbons E, Dalkin S, Wright J, Valderas J, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. Journal of Patient-Reported Outcomes. 2018;2:42. https://doi.org/10.1186/s41687-018-0061-6

[5]    Morel T, Cano SJ. Measuring what matters to rare disease patients – reflections on the work by the IRDiRC taskforce on patient-centered outcome measures. Orphanet Journal of Rare Diseases. 2017;12:171. https://doi.org/10.1186/s13023-017-0718-x

[6] Cano SJ, Pendrill LR, Melin J, Fisher WP. Towards consensus measurement standards for patient-centered outcomes. Measurement. 2019;141:62-9. https://doi.org/10.1016/j.measurement.2019.03.056

[7] Jones SMW, Gaffney A, Unger JM. Using patient-reported outcomes in measurement-based care: Perceived barriers and benefits in oncologists and mental health providers. J Public Health (Berl). 2021. https://doi.org/10.1007/s10389-021-01580-4

[8] Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, et al. Patient-reported outcome measures in arthroplasty registries. Acta Orthop. 2016;87:9-23. https://doi.org/10.1080/17453674.2016.1181816

[9] van der Willik EM, Hemmelder MH, Bart HAJ, van Ittersum FJ, Hoogendijk-van den Akker JM, Bos WJW, et al. Routinely measuring symptom burden and health-related quality of life in dialysis patients: First results from the Dutch registry of patient-reported outcome measures. Clinical Kidney Journal. 2021;14:1535-44. https://doi.org/10.1093/ckj/sfz192

[10] Cano SJ, Hobart JC. Watch out, watch out, the FDA are about. Developmental Medicine & Child Neurology. 2008;50:408-9. https://doi.org/10.1111/j.1469-8749.2008.00408.x

[11] Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. Frontiers in Public Health. 2018;6:149. https://doi.org/10.3389/fpubh.2018.00149

[12] Krogsgaard MR, Brodersen J, Christensen KB, Siersma V, Kreiner S, Jensen J, et al. What is a PROM and why do we need it? Scandinavian Journal of Medicine & Science in Sports. 2021;31:967-71. https://doi.org/10.1111/sms.13892

[13] Messina S, Frongia AL, Antonaci L, Pera MC, Coratti G, Pane M, et al. A critical review of patient and parent caregiver oriented tools to assess health-related quality of life, activity of daily living and caregiver burden in spinal muscular atrophy. Neuromuscular Disorders. 2019;29:940-50. https://doi.org/10.1016/j.nmd.2019.10.001

[14] Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27:1147-57. https://doi.org/10.1007/s11136-018-1798-3

[15] Slade A, Isa F, Kyte D, Pankhurst T, Kerecuk L, Ferguson J, et al. Patient reported outcome measures in rare diseases: A narrative review. Orphanet Journal of Rare Diseases. 2018;13:61. https://doi.org/10.1186/s13023-018-0810-x

[16] Comins JD, Brodersen J, Siersma V, Jensen J, Hansen CF, Krogsgaard MR. How to develop a condition-specific PROM. Scandinavian Journal of Medicine & Science in Sports. 2021;31:1216-24. https://doi.org/10.1111/sms.13868

[17] Verhaart IEC, Robertson A, Wilson IJ, Aartsma-Rus A, Cameron S, Jones CC, et al. Prevalence, incidence and carrier frequency of 5q–linked spinal muscular atrophy – a literature review. Orphanet J Rare Dis. 2017;12:124. https://doi.org/10.1186/s13023-017-0671-8

[18] Sugarman EA, Nagan N, Zhu H, Akmaev VR, Zhou Z, Rohlfs EM, et al. Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: Clinical laboratory analysis of>72,400 specimens. Eur J Hum Genet. 2012;20:27-32. https://doi.org/10.1038/ejhg.2011.134

[19] Russman BS. Spinal Muscular Atrophy: Clinical Classification and Disease Heterogeneity. J Child Neurol. 2007;22:946-51. https://doi.org/10.1177/0883073807305673

[20] Farrar MA, Kiernan MC. The Genetics of Spinal Muscular Atrophy: Progress and Challenges. Neurotherapeutics; New York. 2015;12:290-302. http://dx.doi.org.proxy.hil.unb.ca/10.1007/s13311-014-0314-x

[21] Farrar MA, Park SB, Vucic S, Carey KA, Turner BJ, Gillingwater TH, et al. Emerging therapies and challenges in spinal muscular atrophy. Annals of Neurology. 2017;81:355-68. https://doi.org/10.1002/ana.24864

[22] Mercuri E, Darras BT, Chiriboga CA, Day JW, Campbell C, Connolly AM, et al. Nusinersen versus Sham Control in Later-Onset Spinal Muscular Atrophy. The New England Journal of Medicine; Boston. 2018;378:625-35. http://dx.doi.org.ezproxy.library.dal.ca/10.1056/NEJMoa1710504

[23] Mercuri E, Pera MC, Scoto M, Finkel R, Muntoni F. Spinal muscular atrophy — insights and challenges in the treatment era. Nature Reviews Neurology. 2020;16:706-15. https://doi.org/10.1038/s41582-020-00413-4

[24] Mercuri E, Sansone V. Nusinersen in adults with spinal muscular atrophy: New challenges. The Lancet Neurology. 2020;19:283-4. https://doi.org/10.1016/S1474-4422(20)30068-5

[25] Slayter J, Hodgkinson VL, Lounsberry J, Brais B, Chapman K, Genge A, et al. A Canadian Adult Spinal Muscular Atrophy Outcome Measures Toolkit: Results of a National Consensus using a Modified Delphi Method. 2020.

[26] Brakemeier S, Stolte B, Thimm A, Kizina K, Totzeck A, Munoz-Rosales J, et al. Assessment of Bulbar Function in Adult Patients with 5q-SMA Type 2 and 3 under Treatment with Nusinersen. Brain Sci. 2021;11:1244. https://doi.org/10.3390/brainsci11091244

[27] Kizina K, Akkaya Y, Jokisch D, Stolte B, Totzeck A, Munoz-Rosales J, et al. Cognitive Impairment in Adult Patients with 5q-Associated Spinal Muscular Atrophy. Brain Sci. 2021;11:1184. https://doi.org/10.3390/brainsci11091184

[28] Osmanovic A, Ranxha G, Kumpe M, Müschen L, Binz C, Wiehler F, et al. Treatment expectations and patient-reported outcomes of nusinersen therapy in adult spinal muscular atrophy. J Neurol. 2020;267:2398-407. https://doi.org/10.1007/s00415-020-09847-8

[29] Osmanovic A, Wieselmann G, Mix L, Siegler HA, Kumpe M, Ranxha G, et al. Cognitive Performance of Patients with Adult 5q-Spinal Muscular Atrophy and with Amyotrophic Lateral Sclerosis. Brain Sciences. 2020;11. http://dx.doi.org.ezproxy.library.dal.ca/10.3390/brainsci11010008

[30] Hagenacker T, Wurster CD, Günther R, Schreiber-Katz O, Osmanovic A, Petri S, et al. Nusinersen in adults with 5q spinal muscular atrophy: A non-interventional, multicentre, observational cohort study. The Lancet Neurology. 2020;19:317-25. https://doi.org/10.1016/S1474-4422(20)30037-5

[31] Coratti G, Cutrona C, Pera MC, Bovis F, Ponzano M, Chieppa F, et al. Motor function in type 2 and 3 SMA patients treated with Nusinersen: A critical review and meta-analysis. Orphanet Journal of Rare Diseases. 2021;16:430. https://doi.org/10.1186/s13023-021-02065-z

[32] Krosschell KJ, Young SD, Cruz R, Mazzella A, Curry M, Peterson I. Best Practices for Physical Therapists & Clinical

Evaluators in Spinal Muscular Atrophy (SMA). CureSMA; 2019.

[33] Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evidence Implementation. 2021;19:3-10. https://doi.org/10.1097/XEB.0000000000000277

[34] Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018;169:467-73. https://doi.org/10.7326/M18-0850

[35] Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evidence Implementation. 2021;19:3-10. https://doi.org/10.1097/XEB.0000000000000277

[36] Roy Rosenzweig Centre for History and New Media. Zotero. 2021.

[37] Litwin M. How to Measure Survey Reliability and Validity. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc.; 1995. https://doi.org/10.4135/9781483348957

[38] Kellar SP, Kelvin EA. Munro's Statistical Methods for Health Care Research. 6th Edition. Wolters Kluwer Health | Lippincott Williams and Wilkins; 2013.

[39] Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: Some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. Int J Integr Care. 2002;2:e15.

[40] Eurich DT, Johnson JA, Reid KJ, Spertus JA. Assessing responsiveness of generic and specific health related quality of life measures in heart failure. Health Qual Life Outcomes. 2006;4:1-14. https://doi.org/10.1186/1477-7525-4-89

[41] Boone WJ. Rasch Analysis for Instrument Development: Why, When, and How? CBE Life Sci Educ. 2016;15. https://doi.org/10.1187/cbe.16-04-0148

[42] Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. MESA Press, 5835S; 1993.

[43] Mayhew AG, James MK, Moore U, Sutherland H, Jacobs M, Feng J, et al. Assessing the Relationship of Patient Reported Outcome Measures With Functional Status in Dysferlinopathy: A Rasch Analysis Approach. Frontiers in Neurology. 2022;13.

[44] Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. Qual Life Res. 2018;27:1159-70. https://doi.org/10.1007/s11136-018-1829-0

[45] Mongiovi P, Dilek N, Garland C, Hunter M, Kissel JT, Luebbe E, et al. Patient Reported Impact of Symptoms in Spinal Muscular Atrophy (PRISM-SMA). Neurology. 2018;91:e1206-14. https://doi.org/10.1212/WNL.0000000000006241

[46] RStudio Team. RStudio: Integrated Development for R 2020.

[47] Wickham H. ggplot2: Elegant Graphics for Data Analysis 2016.

[48] R Core Team. R: A language and Environment for Statistical Computing 2021.

[49] Pedersen TL, RStudio. ggraph: An Implementation of Grammar of Graphics for Graphs and Networks 2021.

[50] Csardi G, Tamas N. The igraph software package for complex network research. 2021.

[51] Gansner ER, Koren Y, North S. Graph Drawing by Stress Majorization. In: Pach J, editor. Graph Drawing, Berlin, Heidelberg: Springer; 2005, pp. 239-50. https://doi.org/10.1007/978-3-540-31843-9_25

[52] Katz L. A new status index derived from sociometric analysis. Psychometrika. 1953;18:39-43. https://doi.org/10.1007/BF02289026

[53] Bonacich P, Lloyd P. Eigenvector-like measures of centrality for asymmetric relations. Social Networks. 2001;23:191-201. https://doi.org/10.1016/S0378-8733(01)00038-7

[54] Elsheikh B, Prior T, Zhang X, Miller R, Kolb SJ, Moore D, et al. An analysis of disease severity based on SMN2 copy number in adults with spinal muscular atrophy. Muscle & Nerve. 2009;40:652-6. https://doi.org/10.1002/mus.21350

[55] Duong T, Wolford C, McDermott MP, Macpherson CE, Pasternak A, Glanzman AM, et al. Nusinersen Treatment in Adults With Spinal Muscular Atrophy. Neurology: Clinical Practice. 2021;11:e317-27. https://doi.org/10.1212/CPJ.0000000000001033

[56] Kissel JT, Elsheikh B, King WM, Freimer M, Scott CB, Kolb SJ, et al. SMA valiant trial: A prospective, double-blind, placebo-controlled trial of valproic acid in ambulatory adults with spinal muscular atrophy. Muscle Nerve. 2014;49:187-92. https://doi.org/10.1002/mus.23904

[57] Recommendations concerning the information required to monitor nusinersen use in real-world settings. Quebec: L'Institut national d'excellence en sante et en service sociaux (INESSS); 2020.

[58] Hansen CF, Jensen J, Brodersen J, Siersma V, Comins JD, Krogsgaard MR. Are adequate PROMs used as outcomes in randomized controlled trials? an analysis of 54 trials. Scandinavian Journal of Medicine & Science in Sports. 2021;31:972-81. https://doi.org/10.1111/sms.13896

[59] Zizzi CE, Luebbe E, Mongiovi P, Hunter M, Dilek N, Garland C, et al. The Spinal Muscular Atrophy Health Index: A novel outcome for measuring how a patient feels and functions. Muscle & Nerve. 2021;63:837-44. https://doi.org/10.1002/mus.27223

[60] Vandervelde L, Van den Bergh PYK, Goemans N, Thonnard J-L. ACTIVLIM: A Rasch-built measure of activity limitations in children and adults with neuromuscular disorders. Neuromuscular Disorders. 2007;17:459-69. https://doi.org/10.1016/j.nmd.2007.02.013

[61] Trundell D, Skalicky A, Staunton H, Hareendran A, Le Scouiller S, Barrett L, et al. Development of the SMA independence scale–upper limb module (SMAIS–ULM): A novel scale for individuals with Type 2 and non-ambulant Type 3 SMA. Journal of the Neurological Sciences. 2022;432:120059. https://doi.org/10.1016/j.jns.2021.120059

[62] Thimm A, Brakemeier S, Kizina K, Munoz Rosales J, Stolte B, Totzeck A, et al. Assessment of Health-Related Quality of Life in Adult Spinal Muscular Atrophy Under Nusinersen Treatment—A Pilot Study. Front Neurol. 2022;12:812063. https://doi.org/10.3389/fneur.2021.812063

[63] Croskerry P. Cognitive forcing strategies in clinical decisionmaking. Annals of Emergency Medicine. 2003;41:110-20. https://doi.org/10.1067/mem.2003.22