Xie et al. Health and Quality of Life Outcomes (2025) 23:52
https://doi.org/10.1186/s12955-025-02386-8

Health and Quality of Life Outcomes

REVIEW

Open Access

Patient reported outcomes in the FDA approved drugs for systemic rheumatic diseases (2013–2024)

Yan Xie[1†], Yang Liu[2,3,4†], Yunhe Qin[5], Geng Yin[6*], Xiaoyuan Chen[2,3*] and Qibing Xie[1*]

Abstract

Background Patient-reported outcomes (PROs) in systemic rheumatic diseases (SRDs) are in the forefront of clinical research. However, a comprehensive evaluation of PROs in pivotal trials supporting SRD drug approval is lacking.

Objective This study aims to systematically characterize the use of PROs in pivotal trials supporting the US Food and Drug Administration (FDA) approval of SRDs treatments and to assess the quality of reporting.

Methods We reviewed the pivotal trials supporting the approval of SRD indications by FDA since July 2013 to assess the use of PROs, including specific PRO measures (PROMs) and types of endpoints designated. Quality of PRO reporting was assessed according to a modified ISOQoL criteria.

Results From July 1st, 2013, to June 30th, 2024, the FDA approved 43 new SRD indications based on 67 pivotal trials, with 58 trials included in the final analysis. PROs served as multiple types of endpoints in most trials. All 58 reviewed trials utilized PROs as secondary or exploratory endpoints. The numbers of trials that employed PROs as components of primary endpoints, co-primary endpoints, and key secondary endpoints, were 47(81.0%), 4(6.9%), 45(77.6%), respectively. Notably, the inclusion of PROs as components of composite primary endpoints or co-primary endpoints (100% vs. 8.3%, $P<0.001$) and key secondary endpoints (93.5% vs. 16.7%, $P<0.001$) were significantly higher in inflammatory arthritis compared to other SRDs. Regarding PROM types, 37 trials (63.8%) reported both generic and disease-specific PROMs, covering a broad range of domains. Quality of PRO reporting, influenced by disease type and the presence of additional PRO reports, was moderate to poor in 45 trials (81.8%). Key reporting elements, such as the PRO hypothesis, mode of PROMs completion, and extent and reasons for missing PRO data, were documented in fewer than 30% of the trials.

Conclusion PROs significantly impact SRDs drug approval decisions, especially for inflammatory arthritis. However, the overall quality of PRO reporting in pivotal trials of SRDs is suboptimal and needs improvement. Our study provides

[†]Yan Xie and Yang Liu contributed equally to this work.

*Correspondence:
Geng Yin
yingeng1975@163.com
Xiaoyuan Chen
cxya02648@mail.tsinghua.edu.cn
Qibing Xie
xieqibing1971@163.com

Full list of author information is available at the end of the article

a comprehensive summary of PRO application in SRDs trials, highlighting the need for strengthening PRO utilization in non-arthritis SRDs and improving PRO reporting quality in future studies.

Keywords Patient reported outcomes, Systemic rheumatic diseases, Drug approval, Reporting quality

Background

Systemic rheumatic diseases (SRDs) are complex, chronic and heterogeneous autoimmune diseases affecting a broad range of organs and tissues. Patients with SRDs experience a range of devastating manifestations, leading to a significant disease burden globally [1]. SRDs are associated with reduced physical function, impaired work productivity, increased mobility and mortality, and decrements in health-related quality of life (HRQoL) [2, 3, 4].

In recent years, treatment options have expanded for some SRDs but remain limited and even sparse for others. Continuous development of new treatment approaches for SRDs is essential and proper disease assessment is pivotal to guide the choice of therapy. The pathogenic complexity and multifaceted nature of SRDs make it challenging to precisely assess treatment efficacy using traditional objective measures alone. Substantial proportion of SRD patients experience disease flares despite having normal level objective measures. Patients with SRDs prioritize other aspects of their conditions, such as pain, fatigue, sleep quality, functioning, and mental health in disease assessment. This highlights the importance of patient reported outcomes (PROs) in SRDs new drug development [4].

Integrating patient voice in medical product development and approval has received global attentions in recent ten years. Patient Focused Drug Development (PFDD) initiative of the US Food and Drug Administration (FDA) was set up in 2012 to encourage medical product sponsors to integrate patient experiences into the whole stage of drug development and evaluation [5]. Subsequently, a series of relevant authorized documents were released, aiming to guide the collection and use of patient-experience data (PED) in clinical trials [6, 7, 8], particularly for the clinical outcome assessments. Among these, PROs offer a valuable and unique means for patients to express their perspectives on disease and treatment. Indeed, PROs are extensively used in trials of SRDs, especially in inflammatory arthritis (IA). The development and application of PROs in SRDs are in the forefront of clinical research and continue to expand [9].

Nevertheless, inherent methodological challenges exist in the data collection, analysis, and interpretation of PRO data [10]. For example, PROs typically involve one or several validated questionnaires and are collected in multiple timepoints. This approach can hamper respondent enthusiasm, increase the risk of data missing, and undermine the reliability of PRO [11]. Hence, improving the transparency and consistency of PRO reporting in clinical trials to address these methodological challenges is of great importance. Several international guidelines have been published, providing standards of PRO reporting for investigators [12, 13, 14]. Despite these guidelines, suboptimal qualities of PRO reporting have been observed in trials of malignancies [15–16]. Reviews regarding the use of PRO measures (PROMs) have been conducted in some types of SRDs [2–3]. However, several important gaps are still existing in PRO reporting of SRD trials. First, comprehensive assessment of PROMs in trials supporting the new treatment approval for SRDs is lacking. Moreover, to our knowledge, there is almost no comprehensive overview on the quality of PRO reporting in trials of SRDs.

Therefore, this study aimed to comprehensively characterize the use of PROs in pivotal trials supporting the approval of SRDs treatments by the FDA over the latest ten years and to assess the quality of PRO reporting in these trials, trying to clarify the significant role of PROs in SRD new drug development, as well as to address the current limitations of PRO application in SRD trials.

Methods

Search strategy and selection criteria

This cross-sectional analysis systematically reviewed all new drugs, including all new molecular entities (NMEs) and biologic license applications (BLAs), approved for treating SRDs by the FDA from July 1, 2013 to June 30, 2024. Scope of SRDs has been defined in our previous study [17]. As previously described, we use a commercial database (Pharmtube, one of China's authoritative information platform) to extract data of all new applications, including the first approval and latter new-efficacy indication or claim. The list was then verified referring to the Center for Drug Evaluation and Research calendar year approval lists from the FDA website. We excluded the approvals pertaining to new combinations or new formulations. Besides, the approvals for treating pediatric patients were also excluded due to the consideration that these patients may be unable to self-report the PRO data precisely. We use Drugs@FDA, the FDA-approved drugs database, to get the FDA medical review document and labels, which were then used to obtain information of pivotal trials. More details of pivotal trials were identified by searching Pubmed, Embase, and the National Institutes of Health (NIH) trials registry, using the following strategy: ("trial number" OR "trial name" OR "drug name") AND ("specified indication"). More details of the search strategy were provided in supplementary materials. First

publications reporting results of primary and key secondary endpoints and subsequent publications specified for PRO reporting were defined as main publications and additional PRO reports of the trial, respectively. The whole screening process was conducted by two researchers (Yan Xie and Yang Liu) independently, and disagreements were mediated by discussion.

Data extraction and quality assessment of PRO reporting

For each pivotal trial, main publications, additional PRO reports, and results posted on the clinical trial website were carefully reviewed by two researchers (Y.X and Y.L) to extract the PRO data. Details regarding the basic drug information, trial design, intention-to-treat (ITT) population, and PROMs (including PROM names, categories, and type of endpoints), were collected from all of the trials that reported PRO information. PROMs' categories and frequencies of usage in SRDs pivotal trials were analyzed and summarized. All of the extraction processes were completed independently by the two reviewers. Discussions were carried out to deal with any disagreements.

Quality of PRO reporting was determined by the degree of adherence to a modified ISOQoL checklist, and only trials with PROs as primary, co-primary or secondary endpoints were included for further evaluation. The original ISOQoL-recommended PRO reporting standards include 17 items for all studies with PROs regardless of whether the PRO is a primary or secondary endpoint [12]. It also provides 11 additional standards for trials where PRO is a primary endpoint. Given that PROs mainly present as composites of primary endpoints or secondary endpoints in SRDs trials, we developed the modified ISOQoL standards based on the framework of the 17 standards. The modified checklist included 13 items, of which 2 items were further subdivided, resulting in 16 items to analyze at last. More details of the modified ISOQoL checklists are presented in supplementary materials.

For each item, the selection of "yes" would be scored as 1 and "no" as 0. The total adherence score (TAS) was defined as the sum of the item scores and the maximum adherence score to the modified ISOQoL standards was 16 for each trial. Higher TAS represents better PRO reporting quality. The degree of adherence to the modified standards was ranked for each trial as follows: good ($TAS \geq 12$), moderate ($8 \leq TAS < 12$), or poor ($TAS < 8$).

Considering that SRDs clinical trials commonly involved multiple PROs endpoints, we simplified the quality assessment by focusing on one or several PROs to evaluate for each trial. The choice of the PROs for assessment depended on the types of endpoints they presented, following this sequence: (1) primary or co-primary endpoints; (2) key secondary endpoints (also present as components of primary endpoints); (3) other secondary endpoints (also present as components of primary endpoints); (4) key secondary endpoints only; (5) other secondary endpoints only. This sequence was based on the importance of the PROs, trying to present the theoretically highest reporting quality of each pivotal trial. All relevant publications of each trial were evaluated collectively. Two reviewers conducted the quality evaluations.

Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Statistics and data analysis

We used descriptive statistics to characterize data. To compare the PRO utilization between trials of IAs and other SRDs, the $\chi 2$ or Fisher exact test was used. Multivariable linear regression model was fitted to estimate the association of higher PRO reporting scores with the following factors: type of investigating drugs (biologics vs. small molecules); type of endpoints the PRO represents (co-primary vs. key secondary vs. others); sample size ($n > 400$ vs. $n \leq 400$); disease type (IAs vs. other SRDs); year of first publication for the prespecified PRO of quality assessment; and additional PRO reports for the prespecified PRO of quality assessment (yes vs. no).

All statistical analyses were two-sided and were done with Prism (version 9.1.1). P-value less than 0.05 was regarded as statistically significant.

Results

Characteristics of pivotal trials included

From July 1st, 2013, to June 30th, 2024, the FDA approved 43 new SRD indications based on 67 pivotal trials. Information of all approved drugs and supporting trials is provided in supplementary Table 1. Among these trials, 7 trials supporting the approval for treating pediatric SRDs and another 2 trials reported no PRO information were excluded from further analysis. Thus, 58 trials, including 12 trials for rheumatoid arthritis (RA), 21 trials for psoriatic arthritis (PsA), 13 trials for spondyloarthrits (SpA), 5 trials for systemic lupus erythematosus, 3 trials for systemic vasculitis, and 4 trials for other SRDs, were identified for the final PRO analysis (Table 1).

Among the 58 trials analyzed, nearly two-thirds of the trials ($n = 36$, 62.1%) investigated biologics, while the remaining studied small molecules. All of these trials were randomized controlled trials. ITT population was larger than 400 in more than half of the trials ($n = 32$, 55.2%). Generally, PROs served as multiple types of endpoints simultaneously in a single SRD trial. We found that all reviewed trials used PROs as secondary or exploratory endpoints ($n = 58$, 100%). Notably, PROs were included as components of primary endpoints in most of the trials

Table 1 Characteristics of SRDspivotal trials with PRO data(*N* = 58, FDA; 2013–2024)

	Number of trials (*N*, %)
Type of investigating drugs	
Biologics	36 (62.1%)
Small molecules	22 (37.9%)
Type of SRDs	
RA	12 (20.7%)
PsA	21 (36.2%)
SpA (including AS and nr-axSpA)	13 (22.4%)
SLE (including LN)	5 (8.6%)
Systemic vasculitis	3 (5.2%)
Others	4 (6.9%)
Type of PRO endpoints	
Co-primary endpoints	4 (6.9%)
Components of composite primary endpoints	47 (81.0%)
Key secondary endpoints	45 (77.6%)
Secondary and/or exploratory endpoints only	10 (17.2%)
Exploratory endpoints only	3(5.2%)
PRO data reporting strategy	
Reported in primary manuscript and additional PRO report	39 (67.2%)
Only reported in primary manuscript	16 (27.6%)
Only reported in additional PRO report	3 (5.2%)
Sample size (ITT population)	
≤ 200	5 (8.6%)
200–400	21 (36.2%)
400–1000	29 (50.0%)
>1000	3 (5.2%)

AS, ankylosing spondylitis; ITT, intention-to-treat; LN, lupus nephritis; nr-axSpA, non-radiographic axial spondyloarthritis; PRO, patient reported outcome; PsA, psoriatic arthritis; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; SpA, spondyloarthritis; SRDs, systemic rheumatic diseases

(*n* = 47, 81.0%), especially for those for IAs (*n* = 46, 79.3%). PROs were also included as co-primary endpoints in 4 of the trials (6.9%, all for treating IAs) and key secondary endpoints in 45 (77.6%) of the trials. Only 10 (17.2%) trials, all of which targeted SRDs other than IAs, reported PROs as solely secondary and/or exploratory endpoints. Among the 10 trials, three included PROs as exploratory endpoints only. Compared with non-IA SRDs, the proportion of trials that included PROs as components of composite primary endpoints or co-primary endpoints(8.3% vs. 100%, $P < 0.001$) and key secondary endpoints(16.7% vs. 93.5%, $P < 0.001$) were significantly higher in IAs(supplementary Table S2). Most PROs were reported both in the primary publication and additional PRO articles (*n* = 39, 67.2%). Sixteen trials (27.6%) reported PRO data only in the primary manuscript, while 3 trials (5.2%) only in subsequent publications.

Summary and classification of proms in SRDs pivotal trials

As shown in Fig. 1 and supplementary Figure S1, our study identified 43 different PROMs in total, including 24 generic measures and 19 disease specific measures. Among the generic measures, the distribution for assessing various domains was as follows: 3 for QoL, 9 for symptoms, 1 for disease activity, 1 for general function, 5 for work productivity, and 5 for other domains. For disease-specific measures, the distribution was: 8 for QoL, 1 for symptoms, 3 for disease activity, 2 for general function, and 5 for work productivity.

The frequency of application for each PROM was also presented in Fig. 1. Thirty-seven trials (63.8%) reported both generic and disease specific PROMs, while the remaining 21 trials (36.2%) used generic PROMs only. The most broadly used generic measures were the 36-Item Short-Form Health Survey (SF-36, *n* = 55), the pain Visual Analogue Scale or Numeric Rating Scale questionnaire (pain VAS/NRS, *n* = 50), and the Patient Global Assessment of Disease Activity (PtGA, *n* = 48). The European Quality of Life-5 Dimensions (EQ-5D, *n* = 29) and Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue, *n* = 38) were also applicated in more than half of the trials. As for disease specific measures, the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI, *n* = 24) and the Bath Ankylosing Spondylitis Functional Index (BASFI, *n* = 13) were the most frequently used.

Subsequently, we assessed the usage of PROM across different SRDs. We founded that the number of PROMs used were 15 for RA, 25 for PsA, 19 for SpA, and 13 for other SRDs. The numbers of measures specified for these diseases were 4, 5, 3, and 5, respectively. As for mean values of PROMs count for each trial, trials of SpA reported the highest average number of PROM count (*n* = 8.6), while SRDs other than IAs reported the lowest average count (*n* = 3.8). More details were presented in Supplementary Table S3.

Level of PRO reporting adherence to the modified ISOQoL checklist

Three trials with PRO as exploratory endpoints only were excluded and 55 trials were involved in this assessment. The overall quality of PRO reporting in SRD trials and that categorized by disease type were summarized in Table 2.

As shown in Fig. 2, the degrees of adherence to the modified standards for each trial varied across trials: good in 10 trials (18.2%), moderated in 28 trials (50.9%), and poor in 17 trials (30.9%). Most trials reported moderate to poor degree of adherence, especially in trials of SRDs other than IAs.

Seven items were reported in majority of the trials, including rationale for PRO instrument choosing (*n* = 46, 83.6%), validity and reliability of PRO instrument (*n* = 39, 70.9%), PRO endpoint status description (*n* = 53, 96.4%), prespecified analysis plan of PRO (*n* = 49, 89.1%), type I

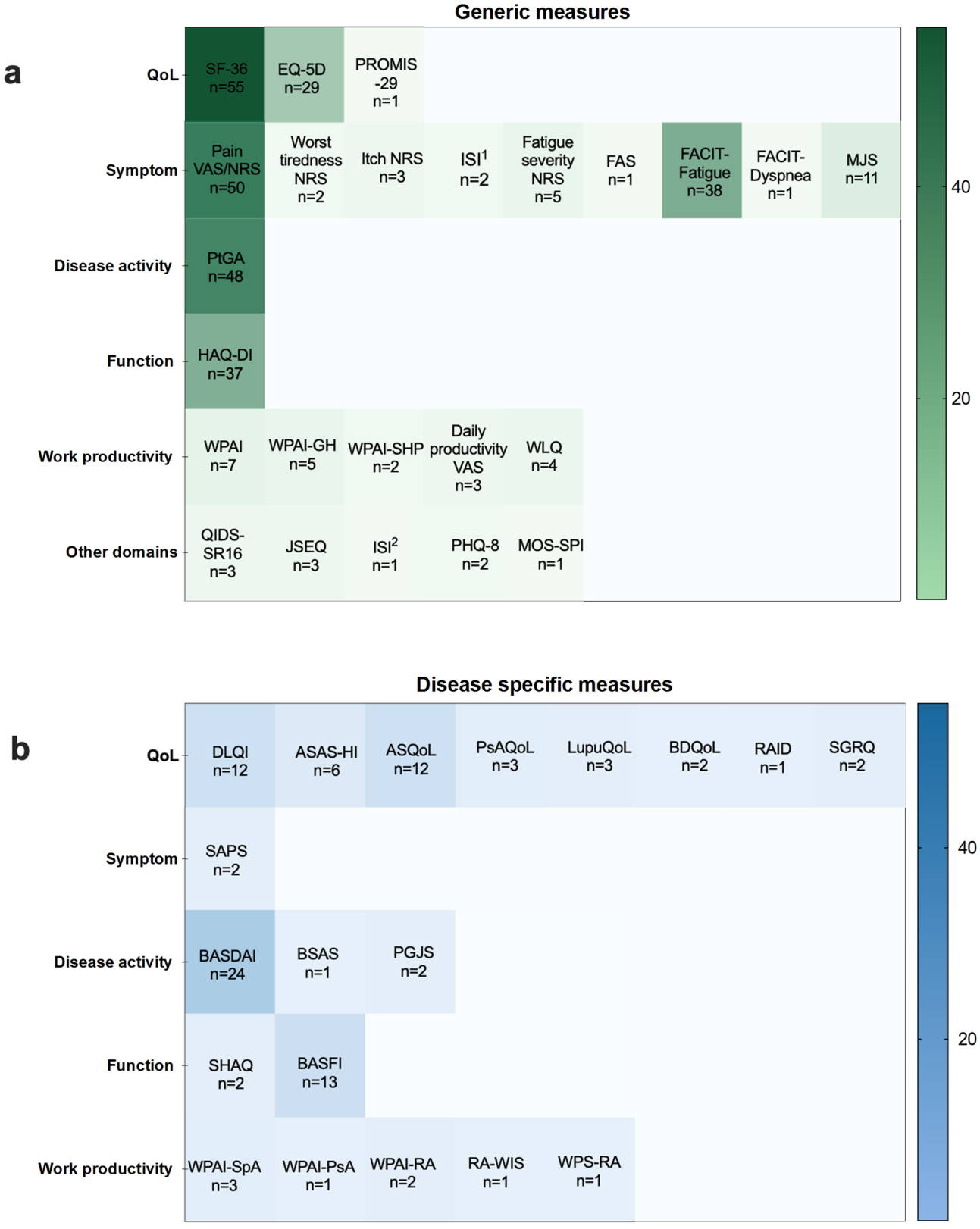

Fig. 1 PROMs in pivotal trials supporting SRD drug approval and overall PRO reporting quality (FDA;2013–2024). Generic (**a**) and disease-specific (**b**) PROMs used in pivotal trials of SRD drugs are summarized by categories. Each PROM was presented as a block in the heat-map. The color scale of the heat-map from light to dark expressed the number of trials used the PROM from low to high. Each PROM name and number of trials using it was noted

Table 2 Level of reporting for each item per the modified ISOQoL checklist on SRDs pivotal trials((*N* = 55, FDA; 2013–2024)

Modified ISOQoL item		RA(*n* = 12)	PsA(*n* = 21)	SpA(*n* = 13)	Other SRDs (*n* = 9)	Total(*n* = 55)
Introduction, background, and objectives						
The PRO hypothesis should be stated and should specify the relevant PRO domain(s) if applicable		3(25.0%)	5(23.8%)	3(23.1%)	1(11.1%)	12(21.8%)
Methods						
Outcome						
The mode of administration of the PRO tool and the methods of collecting data (e.g., telephone, other) should be described		3(25.0%)	5(23.8%)	4(30.8%)	3(33.3%)	15(27.3%)
The rationale for choice of the PRO instrument used should be provided		11(91.7%)	19(90.5%)	10(76.9%)	6(66.7%)	46(83.6%)
Evidence of PRO instrument validity and reliability should be provided or cited		8(66.7%)	17(90.0%)	9(69.2%)	5(55.6%)	39(70.9%)
The intended HRQL data collection schedule should be provided		10(83.3%)	8(38.1%)	8(61.5%)	6(66.7%)	32(58.2%)
PROs should be identified in the trial protocol; post hoc analyses should be identified		3(25.0%)	9(42.9%)	8(61.5%)	5(55.6%)	25(45.5%)
The status of PRO as either a primary or secondary outcome should be stated		11(91.7%)	21(100%)	13(100%)	8(88.9%)	53(96.4%)
Statistical methods						
*There should be evidence of appropriate statistical analysis and tests of statistical significance for each PRO hypothesis tested	◆ Prespecified analysis plan	11(91.7%)	20(95.2%)	13(100%)	5(55.6%)	49(89.1%)
	◆ Type I error control for multiplicity	9(75.0%)	17(90.0%)	13(100%)	3(33.3%)	42(76.4%)
	◆ Prespecified MCID	11(91.7%)	15(71.4%)	3(23.1%)	3(33.3%)	32(58.2%)
Statistical approaches for dealing with missing data should be explicitly stated		11(91.7%)	21(100%)	13(100%)	5(55.5%)	50(90.9%)
The extent of missing data should be stated		1(8.3%)	3(14.3%)	0	1(11.1%)	5(9.1%)
Results						
Participant flow						
A flow diagram or a description of the allocation of participants and those lost to follow-up should be provided for PROs specifically		0	0	0	0	0
The reasons for missing data should be explained		0	2(9.5%)	0	0	2(3.6%)
Baseline data						
The study patients' characteristics should be described, including baseline PRO scores		12(100%)	21(100%)	12(92.3%)	6(66.7%)	51(92.7%)
Discussion						
Strength and limitations related to PRO of the trial should be discussed		10(83.3%)	5(23.8%)	6(46.2%)	2(22.2%)	23(41.8%)

Data are shown as n (%). Items listed in this table were originated from the ISOQoL checklist that applicable to all randomized controlled trials, and were modified based on the disease specificity of SRDs and analysis purpose of this study

*The item was divided into three subitems and adherence to each subitem was scored as 1

HRQL, health-related quality of life; ISOQoL, International Society of Quality of Life Research; MCID, minimal clinically important difference; PRO, patient reported outcome; PsA, psoriatic arthritis; RA, rheumatoid arthritis; SpA, spondyloarthritis; SRDs, systemic rheumatic diseases

error control for multiplicity($n = 42$, 76.4%), statistical approaches for dealing with missing data($n = 50$, 90.9%), and baseline assessment of PRO ($n = 51$, 92.7%). Five items were documented in less than 30% of trials, indicating very poor reporting. Among them, the item of PRO hypothesis and relevant PRO domains was reported in 12 trials (21.8%). Details of the mode of PRO completing and the methods of data collection were documented on 15 trials (27.3%). Two items regarding the extent and reasons for missing data were reported in 5(9.1%) and 2(3.6%) trials, respectively. No trial reported any information on participant allocation and follow up specific to the PROs. The rest 4 items were reported in 41.8-58.2% of the trials, with moderate to poor degree of reporting. Information on intended schedule of PRO collection and prespecified MCID were both identified in 32 trials (58.2%). Twenty-five trials (45.5%) provided available protocol for PROs. The strength and/or limitation related to PRO were discussed in 23 trials (41.8%).

Multivariate analysis of factors associated with greater completeness of PRO reporting

Multivariable linear regression analysis (Table 3) identified two independent factors significantly affecting the quality of PRO reporting, including disease type (β=-5.246, $P = 0.02$) and the presence of additional PRO reports ($\beta = 2.382$, $P = 0.004$). Trials of IAs showed significantly higher completeness of PRO reporting than trials

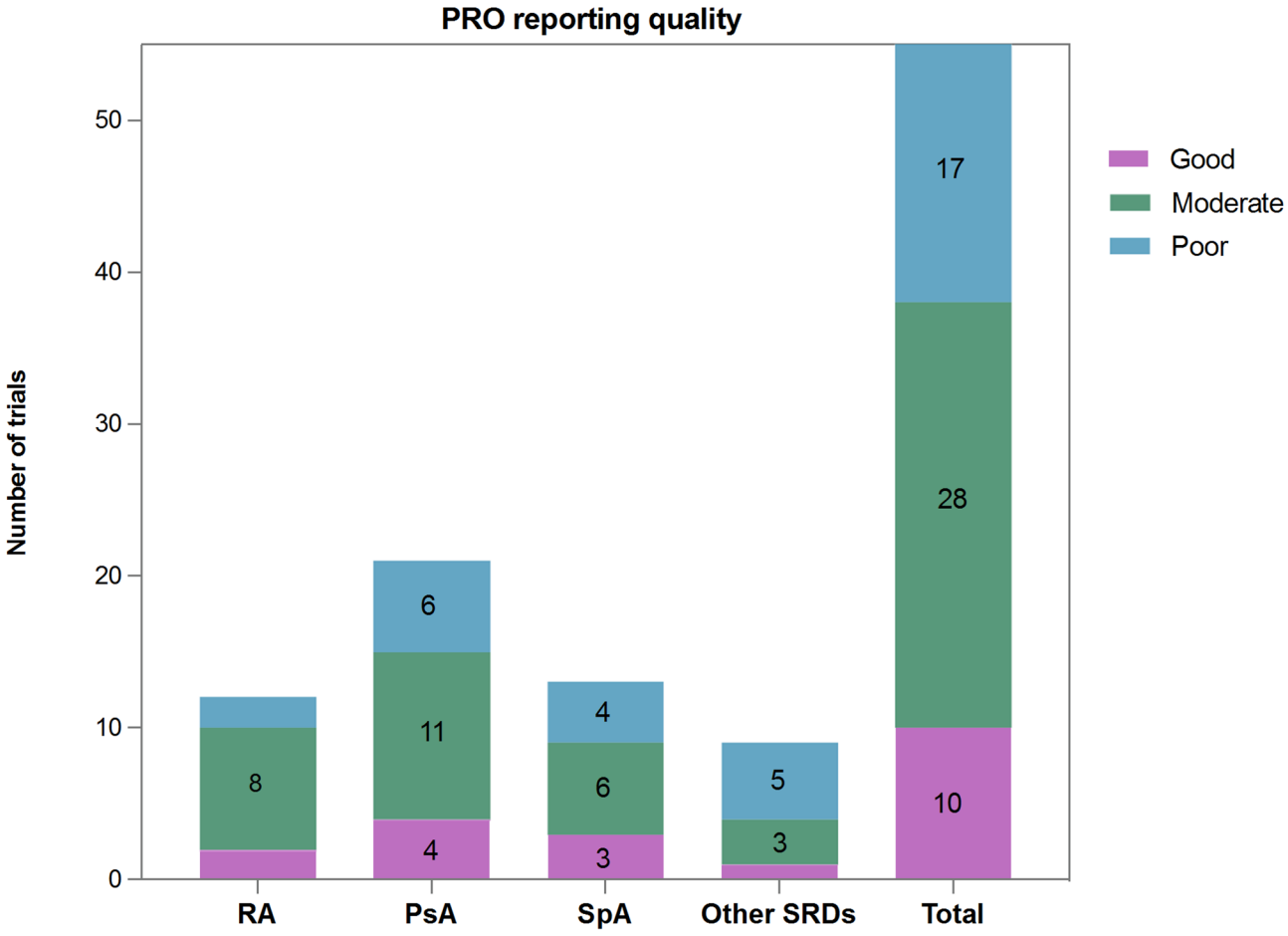

Fig. 2 PRO reporting quality ranking in trials of different SRD indications (FDA; 2013–2024). Numbers of trials for different categories of PRO reporting quality ranking were noted on the middle of the column. PsA, psoriatic arthritis; RA, rheumatoid arthritis; SpA, spondyloarthritis; SRDs, systemic rheumatic diseases

Table 3 Multivariate analysis of factors associated with greater completeness of PRO reporting

	β(95%CI)	*P* value
Drug type (biologics vs. small molecules)	-0.944(-2.460 to 0.528)	0.217
Type of endpoints (co-primary vs. key secondary vs. others)	0.855(-1.064 to 2.771)	0.375
Sample size($n>400$ vs. $n\leq 400$)	-0.600(-2.397 to 1.197)	0.505
Disease type (IAs vs. other SRDs)	-5.246(-9.609 to -0.882)	0.020
Year of first publication for the prespecified PRO of quality assessment	0.127(-0.158 to 0.412)	0.376
Additional PRO reports for the prespecified PRO of quality assessment (yes vs. no)	2.382(0.826 to 3.939)	0.004

CI, coefficient interval; IAs, inflammatory arthritis; PRO, patient reported outcome; SRDs, systemic rheumatic diseases

of other SRDs. The presence of additional reports on the prespecified PRO for quality assessment also indicated better quality of reporting.

Discussion

This comprehensive review showed that PROs were reported in majority of the trials supporting the approval of SRDs treatments by the FDA. Nearly two-thirds of the trials reported both generic and disease specific PROMs, covering wide range of domains. However, the quality of PRO reporting was generally suboptimal (categorized as moderate or poor), influenced by disease type and the presence of additional PRO reports. Trials for IAs and those with additional PRO publications demonstrated better completeness in PRO reporting.

Our findings highlight the essential position of PROs in SRDs assessment, which contrast sharply with their use in other fields. For instances, in oncological drug trials, although included in 67.2% of approvals, PROs mainly served as secondary or exploratory endpoints [18]. Similarly, in diabetes research, only 17% of the trials included

PROs as either primary or secondary outcome [19]. In SRD drug trials, however, our findings demonstrate that PROs served as both key secondary endpoints and components of primary endpoints in over 75% of the trials, underscoring their direct and important role in supporting regulatory approval decision. Additionally, in four trials, PROs were used as co-primary endpoints. Beyond their use in primary and key secondary endpoints, PROs were also extensively utilized as secondary outcomes concurrently to provide additional supportive evidence. The trend of employing PROs as either component of primary endpoints, co-primary endpoints, or key secondary endpoints is obviously more evident in trials of IAs.

Despite their extensive use in SRD trials, PROs were not positioned as the sole primary endpoints. SRD trials always use composite measures which involved PROs as components as the primary endpoints. The subjective nature of PROs makes them susceptible to various factors, including internal factors such as mood, expectations, time, and sentiments, and external factors like treatment, interaction with the health providers, and socioeconomic status [20]. Consequently, PROs are not recommended as the sole primary outcomes and should be supported by objective or functional outcomes in open label settings [20] to reflect patient's condition more accurately.

Our study indicates a high adoption rates of core outcome sets in pivotal trials of SRDs, which can increase the consistency across clinical trials, reduces selective-reporting bias and increase the likelihood that trials will measure relevant and important outcomes [21]. The 43 PROMs identified covered a wide range of domains, with the most common being pain, function, disease activity, and QoL, all included in the core domain sets for SRDs [22]. PROMs for pain, function, and disease activity were typically components of primary outcomes supporting labeling claims, while QoL measures were used as additional supportive evidences. QoL measures, being multidomain, can overlap with pain, function and disease activity, causing redundancy and multiplicity [22]. This multi-domain nature also means QoL items can be affected by variable factors. Regarding the categories of PROMs, we observed a minor difference between the types of generic and specific measures, but more SRD trials employed generic measures. Generic PROMs lack sensitivity but allow for comparison across different conditions and are easy to use, while specific PROMs have greater content validity, credibility, and responsiveness to a specific disease [23]. Although it was generally recommended to use both types of PROMs for robustness results [23], more than one-third of the trials in our study only used generic measures. This is maybe due to limited choices of disease specific PROMs, highlighting the need for future studies to develop new validated disease-specific measures.

Given the significant role of PROs in the SRDs new drug decisions, authoritative agencies should rigorously evaluate the reliability of PROs when assessing a new drug's efficacy. This further emphasizes the importance of high-quality PRO reporting. Despite the extensive use of PROs in SRDs, the reporting of PROs is still suboptimal in pivotal trials that granted the SRD drug's approval, especially in the following methodological aspects: the report of PRO hypothesis, mode of PROMs completion, and extent and reasons for missing PRO data. The results provide supports to a recent published study by Manswell et al. [24], which revealed poor adherences of SRD trials to Consolidated Standards of Reporting Trials (CONSORT)-PRO standards. In the study, nearly 70% of the trials met only two or less items of the standards. Notably, these methodological flaws in PRO reporting are not unique to SRD trials but are similarly prevalent in some other therapeutic areas (e.g., oncology, diabetes, and alcohol use disorder), revealing a common challenge in PRO reporting transcending different disease specialties [15, 16, 19, 25]. Guidelines emphasize the importance of PRO-specific hypothesis [12, 13, 14], yet only 22% of the trials prespecified the PRO hypothesis. Kyte et al. reported similar results in a systematic review of PRO-specific content in trial protocols of different areas, with less than 20% reported PRO related hypothesis [26]. PROMs can be multidimensional or unidimensional and are generally assessed at several timepoints. The absence of a prespecified hypothesis increased the risk of multiple statistical testing and selective reporting, thus hampering the robustness of PRO results [12].

Our study also revealed weaknesses on reporting the PRO data administration in SRDs trials. PROs should be self-reported without any external interpretation4. Discrepancies between the patient self-reported outcomes and proxy- or observer-completed outcomes have been reported [27, 28], highlight the need for clear identification of true responders to avoid potential bias. Likewise, PRO data can be collected by various methods, each with its strengths and limitations [29]. For example, paper-based collection method is suitable for the elderly, but prone to data entry errors [29]. Omitting the mode of PRO data collection prevent the readers from assessing potential bias and the respondent burden, which can contribute to data missing [30].

Missing PRO data is a persisting challenge in clinical trials. It can reduce study power and precision, and complicate data interpretation [31]. The degree of bias introduced by missing data depends on multiple factors, including the extent of missing data within the study and each trial arm, the appropriateness of the assumptions of underlying missing mechanism and the subsequent

analysis strategies [31]. Transparent reporting of these factors is crucial for clinicians to assess potential bias in the PRO results. Besides, reasons for missing PRO data should also be provided as the FDA guidance recommended [32]. Although only less than 10% of the trials in our study reported the extent and reasons for missing data, more than 90% stated statistical approaches to handle it. This percentage is much higher than that in oncology trials [15–16]. However, incorrect assumption of the missing data mechanism would lead to inappropriate statistical methods and biased and misleading results [11]. Therefore, beyond statistical handing, a comprehensive strategy should include transparent reporting of the extent and reasons for missing to evaluate the reliability of the PRO results.

Regarding the variables associated with the quality of PRO reporting, our study showed that IA trials have a higher reporting quality than other SRDs. Additionally, the number and types of PROMs in IAs were higher than in other SRDs, indicating better development of PROs in IAs. Laboratory measures, such as erythrocyte sedimentation rate and C-reactive protein, are nonspecific and often show poor correlation with disease activity in IAs. Moreover, a hallmark symptom of IAs like arthralgia is inherently subjective and can only be accurately reported by patients themselves. Consequently, PROs are consistently recommended as parts of composite measures to supplement clinical and laboratory assessments of disease activity [33, 34]. In contrast, some other SRDs, such as lupus, tend to rely more on objective laboratory indices to reflect the disease activity [35]. This is consistent with our finding that a significant higher proportion of IA trials incorporated PROs as components of composite primary endpoints, co-primary endpoints, and key secondary endpoints. Another factor is the risk of selection bias, as most new approvals in our study were for IAs. We also found that publication of a subsequent PRO report indicated a higher PRO reporting quality, consistent with a previous analysis in oncology trials [15]. This is not surprising, as additional reports provide more detailed information.

Several limitations in our study should be noted. First, we only included the pivotal trials associated with the FDA approvals in recent ten years, which may introduce selection bias and restrict a comprehensive review on PROs in SRDs. This can further limit the applicability of our findings to non-FDA-approved indications and reminds the need of relevant analysis in non-pivotal trials of SRDs in the future. Besides, the new approvals during this period were mainly for IAs, limiting the analysis on PROs of other SRDs. Second, despite efforts to minimize subjectivity, the assessment of PRO reporting quality can still be subjective. Third, since the indications were approved between 2013 and 2024, and their supporting pivotal trials were often published earlier, the trials may not reflect current PRO reporting guidelines. Fourth, potential reporting bias should also be acknowledged as trials with non-significant PRO results may have been less likely to publish. Nonetheless, highlighting these drawbacks in PRO reporting provide a benchmark for future assessment and identify opportunities for improvement.

In conclusion, PROs play a significant role in influencing the decisions of SRDs drug marketing, especially in determining the new drug approval of IAs. However, the overall quality of PRO reporting in pivotal trials of SRDs remains suboptimal. Several specific areas, including the report of PRO hypothesis and mode of PROMs completion, as well as the report of missing data, were reported very poorly and need to be improved. These findings highlight the need to strengthen the utilization of PROs in non-IA SRDs and to improve the PRO reporting quality in future SRD trials. With the growing availability of international guidelines and recommendations for PRO analysis, several strategies can be adopted to facilitate these improvements. These include exploring novel PRO measures, developing PRO reporting guidelines, and expanding the utilization of PROs in real-word studies. Such improvements are essential to provide more valuable and comprehensive insights into the overall treatment efficacy of new drugs in future trials.

Abbreviations

ASAS-HI	Assessment of SpondyloArthritis international Society Health Index
ASQoL	Ankylosing Spondylitis Quality of Life
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
BDQoL	Behçet's Disease Quality of Life score
BSAS	Behçet's Syndrome Activity Scale score
DLQI	Dermatology Life Quality Index
EQ-5D	European Quality of Life-5 Dimensions
FACIT	Functional Assessment of Chronic Illness Therapy
FAS	Fatigue Assessment Scale
HAQ-DI	Health Assessment Questionnaire Disability Index
ISI[1]	Itch Severity Item
ISI[2]	Insomnia Severity Index
JSEQ	Jenkins Sleep Evaluation Questionnaire
LupusQoL	Lupus Quality of Life
MJS	Moring Joint Stiffness
MOS-SPI	Sleep Problems Index II domain of the Medical Outcomes Study Sleep Scale
NRS	Numeric Rating Scale
PGJS	Patient Global Joint and Skin Assessment
PHQ-8	Eight-item Patient Health Questionnaire depression scale
PROMs	Patient reported outcome measures
PROMIS-29	Patient-Reported outcomes Measurement Information System-29
PtGA	Patient Global Assessment of Disease Activity
PsAQoL	Psoriatic Arthritis Quality of Life
QIDS-SR16	Quick Inventory of Depressive Symptomatology-Self Reported 16 items
QoL	Quality of Life
RA-WIS	Work Instability Scale for Rheumatoid Arthritis
RAID	Rheumatoid Arthritis Impact of Disease
SAPS	Self-Assessment of Psoriasis symptoms
SF-36	36-Item Short-Form Health Survey

SGRQ	Saint George's Respiratory Questionnaire
SHAQ	Scleroderma Health Assessment Questionnaire
SpA	spondyloarthritis
SRDs	systemic rheumatic diseases
VAS	Visual Analogue Scale
WLQ	Work Limitations Questionnaire
WPAI	Work Productivity and Activity Impairment questionnaire
WPAI-GH	Work Productivity and Activity Impairment Questionnaire-General Health
WPAI-PsA	Work Productivity Activity Impairment-Psoriatic Arthritis
WPAI-RA	Work Productivity Activity Impairment-Rheumatoid Arthritis
WPAI-SHP	Work Productivity Activity Impairment Questionnaire-Specific Health Problem
WPAI-SpA	Work Productivity Activity Impairment-Spondyloarthritis
WPS-RA	Work Productivity Survey-Rheumatoid arthritis

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12955-025-02386-8.

Supplementary Material 1

Supplementary Material 2

Author contributions

Drs Y. Xie and Liu had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Drs Y. Xie and Liu contributed to the article equally. Concept and design: Y.Xie, Liu, Q. Xie. Acquisition, analysis, or interpretation of data: Liu, Y. Xie, Qin. Manuscript drafting and statistical analysis: Y. Xie, Liu. Critical review of the manuscript: All authors. Administrative, technical, or material support: Y. Xie, Qin. Supervision: Q. Xie, Chen, Yin.

Funding

This article was funded by National Natural Science Foundation of China (grant No.82302029) and Sichuan Science and Technology Program (grant No. 24ZDYF0224).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Competing interests

Qin YH is a staff of Pharmcube. The other authors declare no completing interests.

Consent for publication

Not applicable.

Author details

[1]Department of Rheumatology and Immunology, West China Hospital, Sichuan University, Chengdu 610042, China
[2]Tsinghua Clinical Research Institute (TCRI), School of Medicine, Tsinghua University, Beijing, China
[3]Office of Clinical Trial Institute, Beijing Tsinghua Changgung Hospital, Beijing 102218, China
[4]PhD Clinical and Translational Sciences, Duke-NUS Medical School -Singapore (Singapore), Singapore 169857, Singapore
[5]Pharmcube (Beijing) Co., Ltd, Beijing 100102, China
[6]Health Management Center, General Practice Medical Center, West China Hospital, Sichuan University, Chengdu 610042, China

Received: 30 October 2024 / Accepted: 21 May 2025
Published online: 28 May 2025

References

1. Bossuyt X, De Langhe E, Borghi MO, et al. Understanding and interpreting antinuclear antibody tests in systemic rheumatic diseases. Nat Rev Rheumatol. 2020;16:715–26.
2. Mahieu M, Yount S, Ramsey-Goldman R. Patient-reported outcomes in systemic lupus erythematosus. Rheum Dis Clin North Am. 2016;42:253–63.
3. Gudu T, Gossec L. Quality of life in psoriatic arthritis. Expert Rev Clin Immunol. 2018;14:405–17.
4. Barton JL, Katz P. The patient experience: patient-Reported outcomes in rheumatology. Rheum Dis Clin North Am. 2016;42:xv–xvi.
5. U.S. Food and Drug Administration. CDER Patient-focused drug development. https://www.fda.gov/drugs/development-approval-process-drugs/cder-patient-focused-drug-development. Accessed April 20, 2024.
6. U.S. Food and Drug Administration. Patient-focused drug development: collecting comprehensive and representative input; 2020. https://www.fda.gov/media/139088/download. Accessed April 20, 2024.
7. U.S. Food and Drug Administration. Patient-focused drug development: methods to identify what is important to patients; 2022. https://www.fda.gov/media/131230/download. Accessed April 20, 2024.
8. U.S. Food and Drug Administration. Patient-focused drug development: selecting, developing, or modifying fit-for purpose clinical outcome assessments; 2022. https://www.fda.gov/media/159500/download. Accessed April 20, 2024.
9. ide Wit MP, Berlo SE, Aanerud GJ, et al. European league against rheumatism recommendations for the inclusion of patient representatives in scientific projects. Ann Rheum Dis. 2011;70:722–6.
10. Kyte D, Ives J, Draper H, et al. Inconsistencies in quality of life data collection in clinical trials: a potential source of bias? Interviews with research nurses and trialists. PLoS ONE. 2013;8:e76625.
11. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res. 2014;23:440–59.
12. Brundage M, Blazeby J, Revicki D, et al. Patient-reported outcomes in randomized clinical trials: development of ISOQOL reporting standards. Qual Life Res. 2013;22:o991161–1175.
13. Calvert M, Blazeby J, Altman DG, et al. Reporting of patient- reported outcomes in randomized trials: the CONSORT PRO extension. JAMA. 2013;309:814–22.
14. Calvert M, Kyte D, Mercieca-Bebber R, et al. Guidelines for inclusion of Patient-Reported outcomes in clinical trial protocols: the SPIRIT-PRO extension. JAMA. 2018;319:483–94.
15. Chakraborty R, Cannella L, Cottone F, et al. Quality of patient-reported outcome reporting in randomised controlled trials of haematological malignancies according to international quality standards: a systematic review. Lancet Haematol. 2020;7:e892–901.
16. Efficace F, Fayers P, Pusic A, et al. Quality of patient-reported outcome reporting across cancer randomized controlled trials according to the CONSORT patient-reported outcome extension: A pooled analysis of 557 trials. Cancer. 2015;121:3335–42.
17. Liu Y, Xie Y, Qin Y, et al. Control groups in RCTs supporting approval of drugs for systemic rheumatic diseases, 2012–2022. JAMA Netw Open. 2023;6:e2344767.
18. Ge C, Guo K, Li Y. Analysis of patient-reported outcomes in the approval of novel oncology drugs in the united States, 2017–2022. EClinicalMedicine. 2023;59:101953.
19. Newman C, Kgosidialwa O, Dervan L, et al. Quality of patient-reported outcome reporting in trials of diabetes in pregnancy: A systematic review. Diabetes Res Clin Pract. 2022;188:109879.
20. Kluzek S, Dean B, Wartolowska KA. Patient-reported outcome measures (PROMs) as proof of treatment efficacy. BMJ Evid Based Med. 2022;27:153–5.
21. Kirkham JJ, Boers M, Tugwell P, et al. Outcome measures in rheumatoid arthritis randomised trials over the last 50 years. Trials. 2013;14:324.
22. van Tuyl LH, Boers M. Patient-reported outcomes in core domain sets for rheumatic diseases. Nat Rev Rheumatol. 2015;11:705–12.
23. Churruca K, Pomare C, Ellis LA, et al. Patient-reported outcome measures (PROMs): A review of generic and condition-specific measures and a discussion of trends and issues. Health Expect. 2021;24:1015–24.
24. Manswell K, Le V, Henry K, et al. The state of Patient-Reported outcome measures in rheumatology. J Rheumatol. 2024;51(9):920–7.
25. Douglas A, Garrett E, Staggs J, et al. Reporting of patient-reported outcomes in trials on alcohol use disorder: a meta-epidemiological study. BMJ Evid Based Med. 2023;28(1):21–9.

26. Kyte D, Duffy H, Fletcher B, et al. Systematic evaluation of the patient-reported outcome (PRO) content of clinical trial protocols. PLoS ONE. 2014;9:e110229.
27. Kaiser T, Herzog P, Voderholzer U, Brakemeier EL. Out of sight, out of Mind?? High discrepancy between observer- and patient-reported outcome after routine inpatient treatment for depression. J Affect Disord. 2022;300:322–5.
28. Oldenburger E, Devlies J, Callens D, et al. Patient-reported outcomes versus proxy-reported outcomes in supportive and palliative care: a summary of recent literature. Curr Opin Support Palliat Care. 2023;17:113–8.
29. Ahern S, Ruseckaite R, Ackerman IN. Collecting patient-reported outcome measures. Intern Med J. 2017;47:1454–7.
30. Aiyegbusi OL, Roydhouse J, Rivera SC, et al. Key considerations to reduce or address respondent burden in patient-reported outcome (PRO) data collection. Nat Commun. 2022;13:6026.
31. Nielsen LK, Mercieca-Bebber R, Möller S et al. Relationship between reasons for intermittent missing patient-reported outcomes data and missing data mechanisms. Qual Life Res. 2024.
32. U.S. Food and Drug Administration. Core Patient-reported outcomes in cancer clinical trials: Guidance for industry; 2024. https://www.fda.gov/media/131230/download. Accessed February 8, 2025.
33. Kay J, Morgacheva O, Messing SP, et al. Clinical disease activity and acute phase reactant levels are discordant among patients with active rheumatoid arthritis: acute phase reactant levels contribute separately to predicting outcome at one year. Arthritis Res Ther. 2014;16(1):R40.
34. Ma MHY, Defranoux N, Li W, et al. A multi-biomarker disease activity score can predict sustained remission in rheumatoid arthritis. Arthritis Res Ther. 2020;22(1):158.
35. Cruciani C, Zen M, Gatto M, et al. Assessment of disease activity and damage in SLE: are we there yet? Best Pract Res Clin Rheumatol. 2023;37(4):101896.

Publisher's note