Chapter 13
Outcome Measures for Chronic Pain

Natalie Winter

Assessment of chronic pain is very complex and challenging due to the subjective and individually varying perception of pain and the numerous factors influencing it. Nevertheless, in the field of research and clinical routine, standardized methods for recording pain-related parameters are mandatory to enable a more profound evaluation of therapy success and the comparability of different therapy strategies. Various questionnaires and rating scales, also called instruments, have therefore been established for the assessment and outcome measures of pain.

In chronic pain, objective parameters like blood-pressure or certain serum parameters have not been described yet, or are insufficiently informative because they do not correlate with individual suffering or complex constructs such as depression or quality of life [1]. In recent years, therefore, the focus has increasingly been set on the patient's perspective in the assessment of therapeutic success, in the field of clinical research and in the evaluation of health care services [2]. To further operationalize patient's reports, the terms 'patient-reported outcomes' (PRO), 'patient-reported outcome measures' (PROM) and patient-reported experience measures (PREM) were established. PROs are any reports coming directly from patients about how they feel in relation to a health condition and its therapy, without interpretation of the patient's responses by a clinician, or anyone else [2, 3]. They include any treatment or outcome evaluation obtained directly from patients through interviews, diaries or other data collection tools such as hand-held devices and web-based forms [3]. The corresponding PROM is the instrument used to measure PROs. They mostly contain standardized, validated questionnaires that are completed by patients to ascertain perceptions of their health status, perceived level of

N. Winter (✉)
Department of Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University of Tuebingen, Tuebingen, Germany
e-mail: natalie.winter@med.uni-tuebingen.de

J. de Castro, Y. El Miedany (eds.), *Advances in Chronic and Neuropathic Pain*, Contemporary Rheumatology, https://doi.org/10.1007/978-3-031-10687-3_13

impairment, disability, and health-related quality of life [4]. The focus of PROMs might be generic (designed for the use in populations with different medical conditions), disease-specific or condition-specific (designed for the use in a particular disease state, condition, intervention or treatment of interest) [1, 4].

PREMs are tools that gather information on patients' views of their experience whilst receiving care. They are an indicator of the quality of patient care, although they do not measure it directly. Questionnaires are most common forms for PREMs. In contrast to PROMs, PREMs focus on the impact of the process of the care on the patient's experience e.g. communication and timeliness of assistance [4].

When developing and using PROMs and PREMs, the following psychometric properties should be considered:

- Validity—the instrument measures what it is supposed to measure [1, 5]
- Content validity—describes the extent to which the measurements of a construct fully capture its content in all its aspects [5]
- Reliability—the instrument assesses the feature of interest reliably and consistently over time
- Sensitivity to change—the instrument detects treatment-induced changes

The authors of IMMPACT (Initiative on Methods, Measurement and Pain Assessment in Clinical Trials) also recommend that the parameters appropriateness, interpretability and availability should be taken into account when developing and applying PROMs [6]. Although PROs and PROMs are increasingly used, verification and validation of test procedures is not performed on a regular basis. In addition, many limitations result from the heterogeneity of the instruments used [1]. One approach to address this problem is to establish core outcome sets (COS) —defined as a minimum set of most critical outcome domains and corresponding instruments [1, 6, 7]. These sets represent the minimum that should be measured and reported in all clinical trials of a specific condition, but COS are also suitable for use in routine care, clinical audit and research other than randomized trials [8–10]. A register for COS for all diseases as well as instructions for creating COS can be found at https://comet-initiative.org/. In the field of chronic pain research, various expert consortia have already developed recommendations for COS depending on disease, interventions and outcome definitions (Table 13.1). As chronic pain has multiple effects on patients, COS should cover several domains depending on the context of question [5–7]:

- Pain Intensity
- Pain Interference and Physical functioning
- Emotional functioning
- Patient reported global rating

For each domain, there are different instruments, each with different application specifications. In the following, the most frequently used tests will be presented.

Table 13.1 Examples of existing core outcome set recommendations in chronic pain conditions (modified from Pogatzki et al. [1])

Core outcome sets—domains	IMMPACT [6, 7]-chronic pain	PedIMMPACT [32]–(acute and) chronic pain	Low back pain [33, 34]-Low back Pain	COMPACT [35]–CRPS
Pain	– NRS – Usage of rescue medicine – VRS if NRS might be problematic	– 3–4 years: Poker Chip Tool – 4–12 years: Faces Pain Scale-Revised – ≥8 years: VAS – Pain diary	– NRS – 1-week interval – RMDQ – Oswestry Disability Index 2.1	*Intensity*: – NRS – PROMIS-29 Profile 2 *Neuropathic Components*: – SF-MPQ2
Physical functioning	– Multidimensional Pain Inventory Interference Scale *or* – Brief Pain Inventory Interference items	– Functional Disability Inventory [36] – <7 years PedsQL	– Oswestry Disability Index 2.1 – RMDQ	+ social participation: – PROMIS-29 profile 2 – EQ-5D-5L
Emotional functioning	– BDI *or* – Profile of mood states	– Children's Depression Inventory – Revised Child Anxiety and Depression Scale (RCADS) –<7 years PedsQL	–	– PROMIS-29 Profile 2 – PROMIS suicidal ideation question *Self-Efficacy*: – Pain Self-efficacy Questionnaire
Symptoms and adverse events	Passive Capture	Active Capture *(further research needed)*	Number of deaths	*Disease Severity*: – CRPS Severity Score – CRPS symptoms question
Patient's global impression/ ratings of change	Patient Global Impression of Change	–	–	Patient Global Impression of Change

(continued)

Table 13.1 (continued)

Core outcome sets—domains	IMMPACT [6, 7]-chronic pain	PedIMMPACT [32]–(acute and) chronic pain	Low back pain [33, 34]-Low back Pain	COMPACT [35]–CRPS
Other domains	*Patient's disposition and acquisition data*: CONSORT guidelines [37]	*Role Functioning*: – School attendance – PedMIDAS (persistent headache) [38] – PedsQL *Sleep*: – Sleep Habits Questionnaire *(further research needed)* *Economic Factors*: Instruments in progress	*Health-related quality of life*: – Short form health survey 12 – 10-item PROMIS Global Health	*Catastrophizing*: Pain Catastrophizing Scale

COMPACT: Core Outcome Measurement for CRPS Clinical Studies; CONSORT: Consolidated Standards of Reporting Trials; CRPS: complex regional pain syndrome; IMMPACT: Initiative on Methods, Measurement and Pain Assessment in Clinical Trials; NRS: Numeric Rating Scale; PedMIDAS: Pediatric Migraine Disability Assessment; PedsQL: Pediatric Quality of Life Inventory; PROMIS: Patient-reported Outcomes Measurement Information System; RMDQ: The Roland & Morris Disability Questionnaire; SF-MPQ2: short form McGill Questionnaire 2; VAS: Visual Analogue Scale; VRS: Verbal Rating Scale

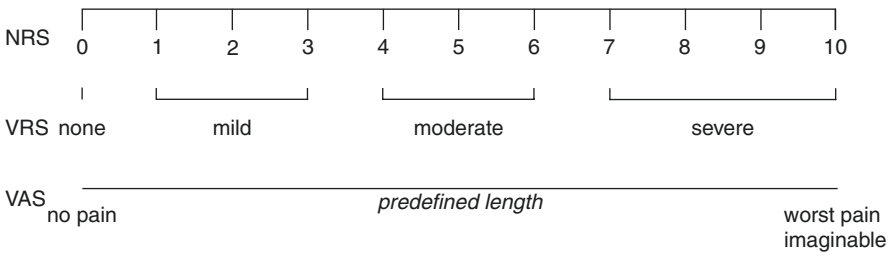

Fig. 13.1 Comparison of Numeric Rating Scale (NRS) , Verbal Rating Scale (VRS) and Visual Analogue Scale (VAS)

Pain Intensity

The most commonly used scales for measuring acute and chronic pain intensity are the **Numeric Rating Scale** (NRS) , **Verbal Rating Scale** (VRS) and **Visual Analogue Scale** (VAS) [11]. The NRS is a unidimensional 11-point scale ranging from 0 points = no pain to 10 points = worst pain (Fig. 13.1). It shows very good correlation with the VAS (correlation ranges from 0.86 to 0.95), in which the patient is asked to mark the pain intensity on a line with a defined length (Fig. 13.1). In terms of sensitivity to changes, both scales are superior to the VRS, which only

consists of four terms describing pain intensity: 'none', 'mild, 'moderate' and 'severe'. In some versions of VRS a fifth category, 'very severe pain', is added. A computerized simulation study of simultaneous recorded data of VAS, NRS and VRS observations documented that the power to detect a difference in pain intensity was higher with the NRS and the VAS data in comparison to the VRS data. Furthermore, the power to detect a difference increases with the magnitude of the difference in pain intensities before and after pain treatment [11, 12]. The IMMPACT authors recommend using VRS as an additional tool to increase comparability between studies [6]. The VRS is also easier to understand, especially for older participants, and quickly completed, so that dropout rates can be reduced if the NRS or VAS are inadequately completed. The VRS and NRS are more suitable for telephone interviews.

The test-retest reliability is very high for NRS for both literate and illiterate people ($r = 0.96$ and 0.95 respectively, studied on patients with rheumatoid arthritis) and VAS (literate $r = 0.94$, illiterate $r = 0.71$, studied on patients with rheumatoid arthritis) [5]. All three scales are freely available.

The **short form McGill Questionnaire 2** (SF-MPQ2) is the specifically extended version of the McGill Questionnaire to include evaluation of neuropathic pain in the assessment of chronic pain. The SF-MPQ2 contains 22 pain descriptors (scored 0–10 using the anchors "none" to "worst possible") on four subscales representing (1) continuous, (2) intermittent, (3) neuropathic, and (4) affective features [13]. For complex regional pain syndrome (CRPS) Dworkin et al. reported good internal consistency for each subscale (ranging from 0.73 to 0.87 across several investigations in large samples). Discriminant validity was supported by significant differences in change scores across a clinical trial in those who considered themselves improved compared with those who did not ($P < 0.002$ for all scales) [14]. Other areas of application for SF-MQ2 include low back pain, painful diabetic neuropathy and cancer pain [13].

Pain Interference and Physical Functioning

Pain Interference refers to all pain-related consequences on relevant aspects of a person's life and include the extent to which pain hinders engagement with social, cognitive, emotional, physical, and recreational activities. Most questionnaires in this area include pain intensity and impact on daily function as well as general functional impairment. **The Roland & Morris Disability Questionnaire** (RMDQ) is a widely used health status measure for low back pain and can be used for clinical purposes or research [5]. The RMDQ contains 24 sentences which describe different limitations of movements or functions for which participants tick those that apply to them that day [5]. The total score ranges from 0 points (no sentences applied) to 24 points (all applied). In several studies psychometric properties have been evaluated: the RMDQ has a good validity and reliability with a range of internal consistency between 0.83 and 0.95 [15, 16], and a range of intraclass correlation coefficients between 0.83 and 0.93 with poorer retest reliability in longer intervals [15, 17–19]. The questionnaire is freely available on the website www.rmdq.org and has already been translated into over 50 languages.

The **Oswestry Disability Index** (ODI) 2.1 has become one of the principal condition-specific outcome measures used in the management of low back pain. The questionnaire is very detailed with 10 categories (pain intensity, personal care, lifting, walking, sitting, standing, sleeping, sex life, social life, traveling), each containing six statements with different degrees of restriction, from which the patient selects the most applicable to him or her. Each category is scored from 0 (no difficulty) to 5 points (maximum difficulty), summed and multiplied by the factor two to obtain a score range from 0 to 100. Test-retest reliability has been shown to be high with intraclass correlation coefficients values ranging from 0.83 to 0.99 [19–23]. Similar to the RMDQ, the reliability decreases with increasing interval between tests. For non-commercial use, the test is free of charge.

As complimentary instruments to disease-specific ones, both the **West Haven-Yale Multidimensional Pain Inventory Interference Scale** (WHYMPI/ MPI) and the **Brief Pain Inventory (BPI) Interference Items** are recommended to measure physical functioning [6]. The WHYMPI is a 52-item, 12-scale inventory that is divided into three parts, which addresses the impact of pain on the patients' lives, the responses of others to the patients' communications of pain, and the extent to which patients participate in common daily activities [24]. Especially the last part is designed to assess physical functioning. The initial study showed good psychometric results: The internal reliability coefficients of all WHYMPI scales ranged from 0.7 to 0.9, the test-retest reliabilities of these scales over a 2-week interval ranged from 0.62 to 0.91 [24].

The BPI was initially designed for the assessment of pain in tumor patients. In the meantime, however, its application in chronic pain associated diseases, such as low back pain, musculoskeletal pain and arthritis, has been widely demonstrated [25–27]. In addition to the MPI, the BPI includes an item to asses pain interference with sleep, which is an important part of the evaluation of physical functioning [6]. A validation study by Tan and his colleagues revealed an acceptable internal consistency (Cronbach alpha coefficients were 0.85 for the intensity items and 0.88 for the interference items), a stable 2-factor structure and responsivity to change [25].

The WHYMPI is freely available. The University of Texas M.D. Anderson Cancer Center holds the copyright, but permission to use the tool can be sought by filling out an online form.

Emotional Functioning

Chronic pain is often accompanied by symptoms of psychological distress and psychiatric disorders, including depression, anxiety and anger [6]. The IMMPACT consensus recommends the **Beck Depression Inventory** (BDI) and the **Profile of Mood Status** (POMS) to assess emotional functioning [6]. Both instruments are very well established in clinical trials as well as in clinical areas such as psychiatry, psychology, cardiology, neurology, obstetrics, nephrology, and others. The BDI consist of 21 items, each item scored from 0 to 3 with cut-offs for the total scores: 0 to 13—minimal depression, 14 to 19—mild depression, 20 to 28 moderate

depression, 29 to 63 severe depression [5]. The POMS, on the other hand, measures six different dimensions of fluctuating mood swings over a certain period of time: 'Depression or Dejection', 'Confusion or Bewilderment', 'Fatigue or Inertia', 'Tension or Anxiety', 'Anger or Hostility', 'Vigor or Activity'. Each feeling is rated by the respondent on a scale from 1 (not at all) to 5 (extremely) to what extent it is currently applicable. Internal consistency is high for both instruments: POMS ranges from 0.63 to 0.96 (Cronbach alpha coefficient) [28], and BDI around 0.9 (Cronbach alpha coefficient) with a test-retest reliability of 0.73 to 0.96 [5].

Both instruments are copyrighted.

Patient Reported Global Rating

To measure global change from the patient's perspective, the Patient Global Impression of Change Scale (PGIC) has been very useful and is recommended by the IMMPACT [6, 7] and COMPACT consortium (Core Outcome Measurement for CRPS Clinical Studies). The PGIC is a 7-point verbal scale to rate the change before and after or under treatment ranging from 1 = no change to 6 = a great deal better. PGIC has shown to correlate significantly with changes in other measurements like Roland Morris Disability Questionnaire, Oswestry Disability Index, EQ-5D and Pain Rating Scale [29]. Test-retest reliability was high (ICC 0.9).

Other Domains

The term 'catastrophizing' was formally introduced by Albert Ellis and subsequently adapted by Aaron Beck to describe a maladaptive cognitive style employed by patients with anxiety and depressive disorders [30]. The pain-related term 'catastrophizing' is broadly conceived as a set of exaggerated and negative cognitive and emotional schemata brought to bear during actual or anticipated painful stimulation [30]. The Pain Catastrophizing Scale (PCS) consists of 13 statements that the participant is asked to rate from 0 = not at all to 4 = all the time, depending on how much these statements apply to the participant when in pain. Internal consistency ranged from 0.86 to 0.95 (Cronbach alpha coefficient) with good validity [31]. The PSC is free for use, for commercial research a payment is required.

PROMIS, Neuro Qol, ASCQ-Me

PROMIS (Patient-reported Outcomes Measurement Information System) is a National Institutes of Health initiative to develop and make self-reported and parent-reported measures of global, physical, mental, and social health for adults and children in the general population and those living with a chronic condition available. Other measurement batteries for specific diseases are also freely accessible: **Neuro QoL** are self-reported and proxy-reported measures of physical, mental, and social health for adults and children living with a neurological condition and **ASCQ-ME** offers self-reported measures of physical, mental, and social health for adults living with sickle cell disease. Many of these measurements have already been tested with regard to their validity and reliability and are also available in different languages. However, the exact specifications should be checked before application. All three databases can be accessed via the Website www.healthmeasures.net

References

1. Pogatzki-Zahn E, Schnabel K, Kaiser U. Patient-reported outcome measures for acute and chronic pain: current knowledge and future directions. Curr Opin Anaesthesiol. 2019;32(5):616–22.
2. Weldring T, Smith SMS. Article commentary: patient-reported outcomes (PROs) and patient-reported outcome measures (PROMs). Heal Serv Insights. 2013;6:61–8.
3. Higgins JP, Green S. Cochrane handbook for systematic reviews. Vol. Version 5. The Cochrane Collaboration. 2008. Version 5.1.0.
4. Kingsley C, Patel S. Patient-reported outcome measures and patient-reported experience measures. BJA Educ. 2017;17(4):137–44.
5. Baranidharan G, Williams A, Wilson S, Cameron P, Tan T. Outcome measures [internet]. British Pain Society. 2019 [cited 2021 May 2]. Available from: https://www.britishpainsociety.org/static/uploads/resources/files/Outcome_Measures_January_2019.pdf
6. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain. 2005;113(1–2):9–19.
7. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain. 2003;106:337–45.
8. Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core outcome set-STAndardised protocol items: the COS-STAP statement. Trials. 2019;20(1):1–7.
9. Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Devane D, et al. Core outcome set–STAndards for reporting: the COS-STAR statement. PLoS Med. 2016;13(10):1–11.
10. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core outcome set-STAndards for development: the COS-STAD recommendations. PLoS Med. 2017;14(11):1–10.
11. Breivik H, Borchgrevink PC, Allen SM, Rosseland LA, Romundstad L, Breivik Hals EK, et al. Assessment of pain. Br J Anaesth. 2008;101(1):17–24.
12. Breivik Hals EK, Björnsson GA, Skovlund E. A comparison of pain rating scales by sampling from clinical trial data. Clin J Pain. 2000;16:22–8.
13. Packham TL, Bean D, Johnson MH, Macdermid JC, Grieve S, Mccabe CS, et al. Measurement properties of the SF-MPQ-2 neuropathic qualities subscale in persons with CRPS: validity , responsiveness , and Rasch analysis. Pain Med. 2018;0(0):1–11.
14. Dworkin RH, Turk DC, Trudeau JJ, Benson C, David M, Katz NP, et al. Validation of the short-form McGill pain questionnaire-2 (SF-MPQ-2) in acute low Back pain. J Pain. 2015;2:1–29.
15. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland- Morris questionnaire. Phys Ther. 2018;76(4):359–65.
16. Spanjer J, Groothoff JW, Brouwer S. Instruments used to assess functional limitations in workers applying for disability benefit: a systematic review. Disabil Rehabil. 2011;33:2143–50.
17. Brouwer S, Kuijer W, Dijkstra PU. Reliability and stability of the Roland Morris disability questionnaire: intra class correlation and limits of agreement. Disabil Rehabil. 2004;26(3):162–5.
18. Jordan K, Dunn KM, Lewis M, Croft P. A minimal clinically important difference was derived for the Roland-Morris disability questionnaire for low back pain. J Clin Epidemiol. 2006;59:45–52.
19. Davidson M, Keating JL. A comparison of five low back disability questionnaires: reliability. Phys Ther. 2002;82(1):8–24.
20. Roland M, Fairbank J. The Roland – Morris disability questionnaire and the Oswestry disability questionnaire. Spine (Phila Pa 1976). 2000;25(24):3115–24.
21. Fairbank JCT, Pynsent PB. The Oswestry disability index. Spine (Phila Pa 1976). 2000;25(22):2940–53.
22. Grotle M, Garratt AM, Krogstad H, Stuge B. Research report reliability and construct validity of. Phys Ther. 2012;92(1):111–23.
23. Miekisiak G, Kollataj M, Dobrogowski J, Kloc W. Validation and cross-cultural adaptation of the. Spine (Phila Pa 1976). 2013;38(4):237–43.

24. Kerns R. The west Haven-Yale multidimensional pain inventory the west Haven-Yale multidimensional inventory (WHYMPI). Pain. 1986;23:345–56.
25. Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the brief pain inventory for chronic nonmalignant pain. J Pain. 2004;5(2):133–7.
26. Keller S, Bann CM, Dodd SL, Schein J, Mendoza TR, Cleeland CS. Validity of the brief pain inventory for use in documenting the outcomes of patients with noncancer pain. Clin J Pain. 2004;20(5):309–18.
27. Celik E, Yalcinkaya E, Atamaz F, Karatas M, Ones K. Validity and reliability of a Turkish brief pain inventory short form when used to evaluate musculoskeletal pain. J Back Musculoskelet Rehabil. 2016;1:1–5.
28. Curran SL, Andrykowski MA, Studts JL. Short form of the profile of mood states (POMS-SF): psychometric information. Psychol Assess. 1995;7(1):80–3.
29. Kamper SJ, Maher CG, Mackay G. Global rating of change scales. J Man Manip Ther. 2009;17(3):163–70.
30. Quartana PJ, M CC, Edwards RR. Pain catastrophizing: a critical review. Expert Rev Neurother. 2009;9(5):745–58.
31. Osman A, Barrios FX, Gutierrez PM, Kopper A, Merrifield T, Grittmann L. The pain catastrophizing scale: further psychometric evaluation with adult samples. J Behav Med. 2000;23(4):351–65.
32. McGrath PJ, Walco GA, Turk DC, Dworkin RH, Brown MT, Davidson K, et al. Core outcome domains and measures for pediatric acute and chronic/recurrent pain clinical trials: PedIMMPACT recommendations. J Pain. 2008;9(9):771–83.
33. Chiarotto A, Boers M, Deyo RA, Al E. Core outcome measurement instruments for clinical trials in non-specific low back pain 1. Pain. 2018;159:481–95.
34. Chiarotto A, Deyo RA, Terwee CB, Boers M, Buchbinder R, Corbin TP, et al. Core outcome domains for clinical trials in non-specific low back pain. Eur Spine J. 2015;24:1127–42.
35. Grieve S, Perez RSGM, Birklein F, Brunner F, Bruehl S, Harden RN, et al. Recommendations for a first core outcome measurement set for complex regional PAin syndrome clinical sTudies (COMPACT). Pain. 2017;158:1083–90.
36. Walker LS, Greene JW. The functional disability inventory: measuring a neglected dimension of child health status. J Pediatr Psychol. 1991;16:39–58.
37. CONSORT [Internet]. [cited 2021 May 2]. Available from: http://www.consort-statement.org/
38. Hershey AD, Powers SW, Vockell AB. Development of a questionnaire to assess disability of. Neurology. 2001;57:2034–9.