

SYSTEMATIC REVIEW

Open Access



How is recovery defined and measured in patients with low back pain? A mixed study systematic review

Michael J. Lukacs^{1,2*} , Nicole Peters¹ , Masakazu Minetama³ , Katie L. Kowalski¹, Meagan Stanley⁴, Keerthana Jayaprakash⁵, David M. Walton¹  and Alison B. Rushton¹ 

Abstract

Background Despite the burden of low back pain (LBP) there is no currently accepted definition for its recovery, nor is there a gold standard for measurement. In addition, it is currently unclear how the perspective of patients are used in making recovery determinations. The purpose of this mixed study systematic review across both quantitative and qualitative literature was to (1) explore how recovery has been defined and measured for patients experiencing LBP, and (2) examine how the perspectives of patients and providers for recovery of LBP align or differ.

Methods This was a mixed study systematic review. Key databases were searched from inception until February 20, 2023: Medline, EMBASE, CINAHL, Cochrane, PEDro looking for sources examining definitions and measures of recovery in patients with LBP. Grey literature was identified through the ProQuest Thesis & Dissertation database. Two reviewers used the Mixed Methods Appraisal Tool for quality assessment of both qualitative and quantitative studies to explore definitions, measurements and perspective of recovery.

Results 466 original studies were included: 12 qualitative studies, 88 quantitative randomized control trials, 348 quantitative non-randomized studies, 16 quantitative descriptive studies, and two mixed methods studies. Most of the time recovery was not defined, with six other themes identified: comparison of scores, in relation to a singular cut-off score, improvement of absence of clinical symptoms, a return to a pre-injury state, change/improvement score from baseline and as a process/trajectory. For recovery measurements, six themes described the data: multiple measures, single measure excluding recovery, a recovery measure, recovery and an additional measure, pain and an additional measure, or indirect/ not specified. Lastly recovery perspectives were made from either the patient, provider, or a combination of patient and provider.

Conclusion For patients living with LBP, the concept of recovery continues to lack consensus for its definition and measurement in patients with LBP. The perspectives of patients were mostly not preserved in making recovery determinations. Urgent action is needed to generate consensus across clinicians, researchers, and patients regarding how recovery should be defined and measured. A multitude of study-specific definitions limit knowledge syntheses and definition of best practice.

*Correspondence:
Michael J. Lukacs
mlukacs2@uwo.ca

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



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Keywords Low back pain, Recovery, Mixed study systematic review

Introduction

Low back pain (LBP) is a leading cause of years lived with disability globally [1] experienced by up to 80% of people during their life [2]. While recovery from LBP is often described as a primary goal of patients, healthcare providers, and funders [3–5], there remains no universally agreed definition for what recovery of LBP constitutes or how it should be identified [6]. In practice, recovery is often measured as a *state* relative to a threshold of measurement rather than a process [7, 8], despite well-known challenges with assigning a single cut-off score to all people [9]. As a result, patients living with pain may feel that the ways in which they are evaluated are not relevant to them and may also not incorporate their own perspectives [10].

Consensus on how best to define, and thus measure recovery has not been achieved across most spinal pain conditions including LBP [4, 6, 11, 12]. Without a unified concept of how to define or measure recovery in patients with LBP, comparison of recovery outcomes across quantitative and qualitative literature can be difficult if not impossible [4]. In 2011, Kamper and colleagues conducted a systematic review to explore the ways recovery had been defined and measured in patients with LBP [6]. While that research synthesized and identified the problem of heterogeneous outcomes, it did not include an examination of how the patient's perspective was integrated into recovery determinations, it excluded surgical literature [6], and is now over a decade old. Inclusion of qualitative research is also needed, due to its ability to preserve the patient's voice [13]. Thus, the aim of this mixed study systematic review (MSSR) is to examine how recovery has been defined and measured for patients with LBP, within both the quantitative and qualitative literature.

Objectives

1. To explore how recovery has been defined and measured for patients experiencing LBP.
2. To examine how the perspectives of patients and providers regarding the recovery of LBP align or differ.

Methods

Design

This is a MSSR reported in line with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) [14] (see S1 File). This review is considered a MSSR as it synthesized the definitions and

measurements of recovery across multiple study types (e.g., both quantitative and qualitative literature sources) using both a systematic search strategy and quality assessment. Additionally, this research also aims to understand how perspectives of patients are integrated in making recovery determinations by including qualitative literature, and to understand if they are in line with those of researchers.

Protocol and Registration

This MSSR was registered in PROSPERO (ID: CRD42022295804). The protocol has been published previously [4] following the referred Reporting Items of Systematic Review and Meta-Analysis Protocols (PRISMA-P) [15, 16]. There were no significant protocol deviations, but the exclusion criteria were clarified to highlight more specifically the population of interest (i.e., musculoskeletal LBP).

Eligibility criteria

Population Patients with musculoskeletal LBP with or without leg pain (i.e., >50% of participants with LBP).

Intervention Either conservative (e.g., physiotherapy or observational) or non-conservative (e.g., surgical) intervention.

Comparator N/A.

Outcome Definitions and/or measurements of recovery. Each study had to report the concept of recovery in either the abstract, methods or results.

Study Design Quantitative, qualitative studies, and observational (e.g., no intervention) of any study design were included. Studies were excluded if they were not written in English and translation was not possible, LBP was non-musculoskeletal (e.g. cancer, infection, metabolic disorders, etc.), or if quality assessment was not possible (e.g., study protocol, opinion pieces, editorial).

Information sources

The Medline, EMBASE, CINAHL, Cochrane, and PEDro databases were searched from inception until February 20, 2023. Grey literature was identified through ProQuest Thesis & Dissertation database. Hand searching of the references of included studies was also performed in duplicate to identify any articles that were inaccurately indexed or missed by the search strategy.

Search Strategy

The search strategy was piloted within MEDLINE (Ovid) and adapted to the other databases (see S2 File). The search strategy used in this study was developed with assistance from a research librarian (MS). All citations were imported into Covidence (Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia, www.covidence.org) at each stage of the review with duplicates removed.

Selection process

Two independent reviewers evaluated the titles and abstracts of articles for exclusion against the eligibility criteria. Full texts of retained manuscripts were then assessed against the eligibility criteria independently by two authors. A third author (AR) with spinal pain research expertise was consulted for any disagreement.

Data Collection process and data items

Data from included studies were extracted by two reviewers using a customized data extraction sheet for all study designs in a duplicate process. Data extracted included: study authors, publication year, country, type of study, study purpose, study participants, classification of LBP (e.g., acute), description of how recovery was defined, measurement of recovery, and perspective of recovery determination (patient vs. provider).

Quality Assessment

For quality assessment, the Mixed Method Appraisal Tool Version 2018 (MMAT) was used enabling use of a single tool for assessment of mixed study research [17]. The MMAT 2018 tool is a quality assessment tool first created in 2006 for multiple study designs including qualitative, randomized control trials (RCTs), quantitative non-randomized, quantitative descriptive and mixed methods studies [17]. The MMAT focuses specifically on reviews that include quantitative, qualitative, and mixed methods studies which was appropriate for this MMSR [18, 19]. If an included study was a secondary analysis of a RCT, or reporting of new results, the quality assessment was informed by the original RCT. In lieu of an established threshold, studies with a MMAT score of 0–2/5 were rated as low quality, 3–4/5 equated to moderate quality, and studies that scored 5/5 were of high quality. This was a deviation from the published protocol. Quality assessment was used to help inform confidence in the findings of the thematic synthesis. Given that the quality of surgical and physiotherapeutic trials has been criticized as being sub-optimal [20, 21], quality assessment of included papers was integral for this work to comment on recommendations for the field going forward. All studies regardless of quality were included in the synthesis due to the nature of this MMSR.

Data synthesis

A convergent qualitative synthesis was used where both quantitative and qualitative literature were synthesized together to answer the central research questions [22, 23]. Specifically, thematic analysis [24, 25] was used to synthesize the included qualitative and quantitative studies to compare definitions, measurements, and perspectives of recovery in patients with LBP. Thematic analysis of the definitions and measurements of recovery took place independently by two of the researchers (MJL and NP) in three steps: line-by-line coding, free-coding, and development of analytical themes [24]. There were no predefined themes, they were generated as interpreted by the researchers for all types of included studies (e.g., quantitative and qualitative). Each initial definition or measurement created the first theme, against which the next definition/measurement was compared and either grouped together with the first or used to create a new one. This process continued until all definition/measurements/perspectives had been categorized with similar others. Next, two reviewers independently examined each individual theme and those that were conceptually similar with another were merged into broader themes. This process proceeded for both measurements and definitions of recovery. Specifically for the perspectives of recovery, definitions and measures were interpreted together to inform the degree to which patient perspectives were preserved. In this way, the stages of analysis moved from specific to general, offering opportunities for interpretation for different applications. For visual presentation, the frequency of themes within definitions, measurements, and perspectives of recovery were then summed and presented as percentages as part of the total synthesis.

Results

Study selection and characteristics

The search strategy identified 27,899 papers. Following exclusion of 7172 duplicates and 19,391 papers from title/abstract review, 1387 studies were identified for full text review. Following exclusions, 500 studies were eligible for inclusion (Fig. 1. PRISMA flow diagram), including 12 qualitative studies (S1 Appendix) 88 quantitative RCTs (S2 Appendix and S3 Appendix), 348 quantitative non-randomized studies (S3 Appendix), 16 quantitative descriptive studies (S4 Appendix), and two mixed methods Studies (S5 Appendix). 34 of the 500 studies were identified as a secondary analysis of a primary RCT or reporting of new results from an original RCT (S6 Appendix). These studies were included in the synthesis but retained the same quality rating as their initial RCT leaving 466 original studies.

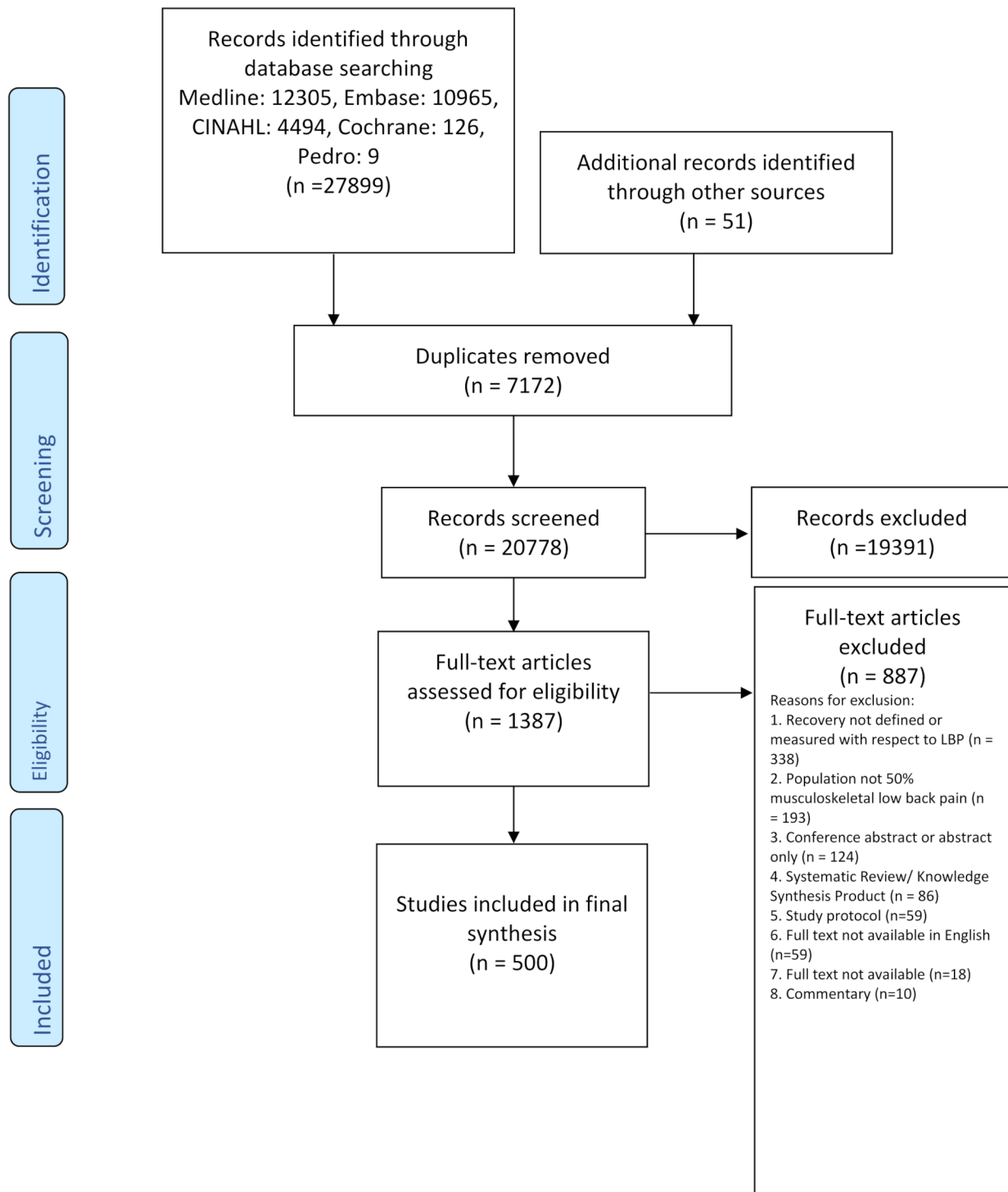


Fig. 1 Flow diagram of literature search

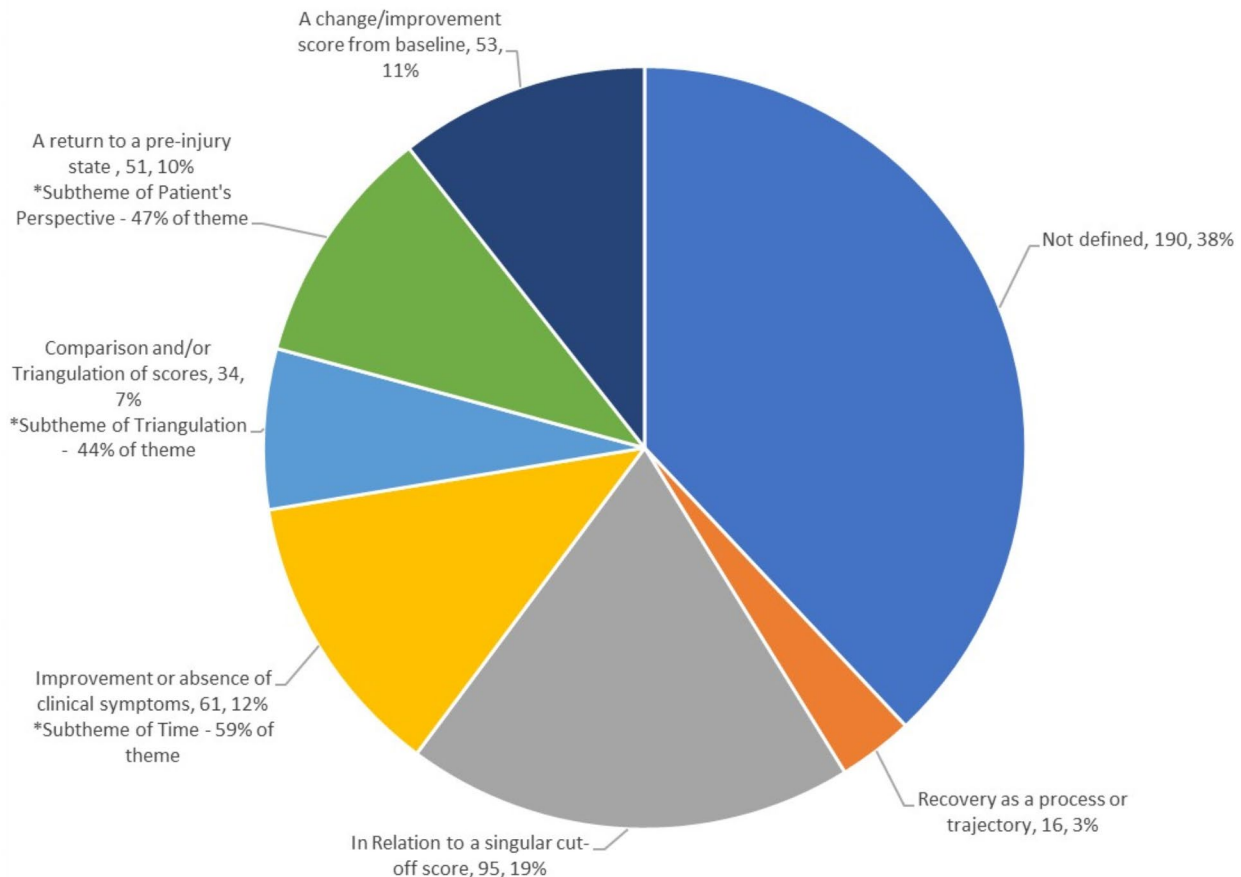
Quality assessment

Most studies were rated as moderate (49%) or low (45.8%) quality. 42% of qualitative studies and 10% of quantitative RCTs were rated as high quality, while five quantitative non-randomized studies were rated as high

quality. See Table 1 for additional detail. Common issues that lead to lower quality varied depending on study design, but examples included issues with randomization or blinding, lack of psychometric support for included

Table 1 Study Design and Quality Assessment (MMAT)

Study Design	Low Quality (n)	Moderate Quality (n)	High Quality (n)	Total (n)
1. Qualitative	1	6	5	12
2. Quantitative Randomized Control (Trials)	40	66	16	122
3. Quantitative non-randomized	179	164	5	348
4. Quantitative Descriptive	8	8	0	16
5. Mixed Methods	1	1	0	2
Total	229	245	26	500

**Fig. 2** Seven main themes of recovery definitions across entire synthesis. Results are presented in both counts and percentage of total synthesis

measures, proxy measures used in place of direct recovery measures, no control for possible co-interventions, or a high loss to follow-up. For qualitative studies, common reasons for lower quality scores pertained to inappropriate qualitative approaches or a lack of coherence between data sources and interpretation.

Definitions of recovery

Thematic analysis identified seven themes describing how recovery had been defined across included studies, three of which had sub-themes.

Recovery not defined Most commonly, recovery was not formally or properly defined with 38% of studies lacking a proper definition (Fig. 2). Typically, studies may have mentioned recovery as an outcome of interest but did not include any formal operational definition. Alternately, studies cited a measure for recovery but failed to indicate its interpretation.

Comparison to a Singular Cut-off Score: The next largest theme for recovery definitions was defined in comparison to a singular cut-off score (19% of studies), whereby patients were deemed recovered by the research team/provider (Fig. 2). Typically, a dichotomous anchor point on a patient-reported outcome measure (PROM) split

patients into recovered and non-recovered categories. Examples included: < 2/10 on the Numeric Pain Rating (NPRS) Scale [26], > six on the Global Perceived Effect/Global Rating of Change (GRC) Scale [27], or score on the Roland-Morris Disability Questionnaire [28] of 0/24.

Absence/Improvement in clinical symptoms Absence/improvement in clinical symptoms was identified from 12% of studies. Examples included terms such as, “*Clinical symptomatic recovery*,” [29] or, “*Symptomatic recovery and/or regression of LDH symptoms*” [30]. This theme had a sub-theme related to time required for absence of clinical symptoms, though the time needed was inconsistent. Examples included, “*Participants needed to report pain no greater than 1/10 during the past week*,” [31] or “*Participants were considered recovered if they had 30 consecutive days with pain no greater than 1 on a 0–10 scale*” [32].

Return to a Pre-injury State In 10% of studies, recovery was defined as a return to a pre-injury state. This took the form of asking patients if they were completely recovered, or questions of whether the patient had returned to normal function. Examples included, “*the total disappearance of low back pain and return to everyday life as it was before the pain*” [33]. This theme also had a sub-theme surrounding incorporation of the perspective of patients. Compared to the other recovery definition themes, this particular theme had many definitions include the patient’s direct perspective. In this way recovery was defined from the perspective of patients as opposed to being labelled as such by study investigators.

Comparison of scores In place of comparison to a singular cut-off score, many studies defined recovery in relation to multiple cut-off scores usually for pain intensity (e.g., NPRS) and disability outcomes (e.g., RMDQ). As a sub-theme approximately 1/3 of included studies used triangulation of outcome scores to deem a patient recovered or not (i.e., three or greater cut-off scores). Across the entire recovery definition synthesis, only 7% of included studies used a comparison/triangulation of scores to define recovery (Fig. 2). Of this theme, almost half of include studies triangulated a recovery definition based on multiple scores. Outside of similarly led investigator groups, these comparisons/triangulations of scores to define recovery were rarely the same. For example, “*fully recovered was defined as scores of one or zero on the Roland-Morris disability scale and five or less on the Oswestry disability scale*” [34].

Change/Improvement from a baseline score For the theme of change/improvement from a baseline score (11% of studies), a dichotomous label was not applied to the measure used, rather recovery was broadly interpreted con-

sidering change from a pre-intervention score. Examples included recovery defined as a ratio related to a baseline score, or whereby high scores broadly equated to better recovery. For example, “*A recovery rate of 75% or greater was regarded as excellent, 50 to 74.9% as good, 25 to 49.9% as fair, and less than 25% as poor*” [35].

Recovery as a process/trajectory The final theme related to recovery defined as a process or trajectory (3% of studies). Definitions under this theme often used a measure such as pain intensity to longitudinally plot the recovery of patients along trajectory patterns. Terms often included the idea of one trajectory representing recovery with the others representing a state of non-recovery related to the presence of moderate/high levels of pain or disability.

Measures of recovery

Six main themes captured how recovery had been previously measured across all included studies (Fig. 3), with four sub-themes.

Multiple measures The most common theme for recovery measurement was through use of multiple measures (35% of studies). This was commonly achieved through pain and disability-focused measures, although motor and sensory function were also commonly measured. This theme included a sub-theme of *inclusion of a recovery construct* or not, with 10% of included studies using multiple measures that included recovery-centric measures (i.e., recovery-direct measure such as GRC or recovery expectations).

Recovery measure The second largest measurement theme related to the use of a *recovery-centric* measurement in isolation, equating to 29% of the thematic synthesis for measures. These measures included the Global Perceived Effect, or Self-Perceived Recovery.

Singular measure excluding recovery 22% of the included studies used a singular measure excluding a recovery-centric measure. This was accomplished through measurement of disability (26% of total theme), pain intensity (48% of total theme) or physical function (26% of total theme) as proxy measures for recovery, each of which were sub-themes. Common measures included using the NPRS, Visual Analog Scale [26], Oswestry Disability Index [36], RMDQ, or description of a return to pre-injury function or work status.

Recovery/Pain plus an additional measure 13% of included measures were categorized as using two measures for measurement of recovery, with either a focus on pain or recovery. Some studies included a recovery-centric measure such as the GRC and a physical function

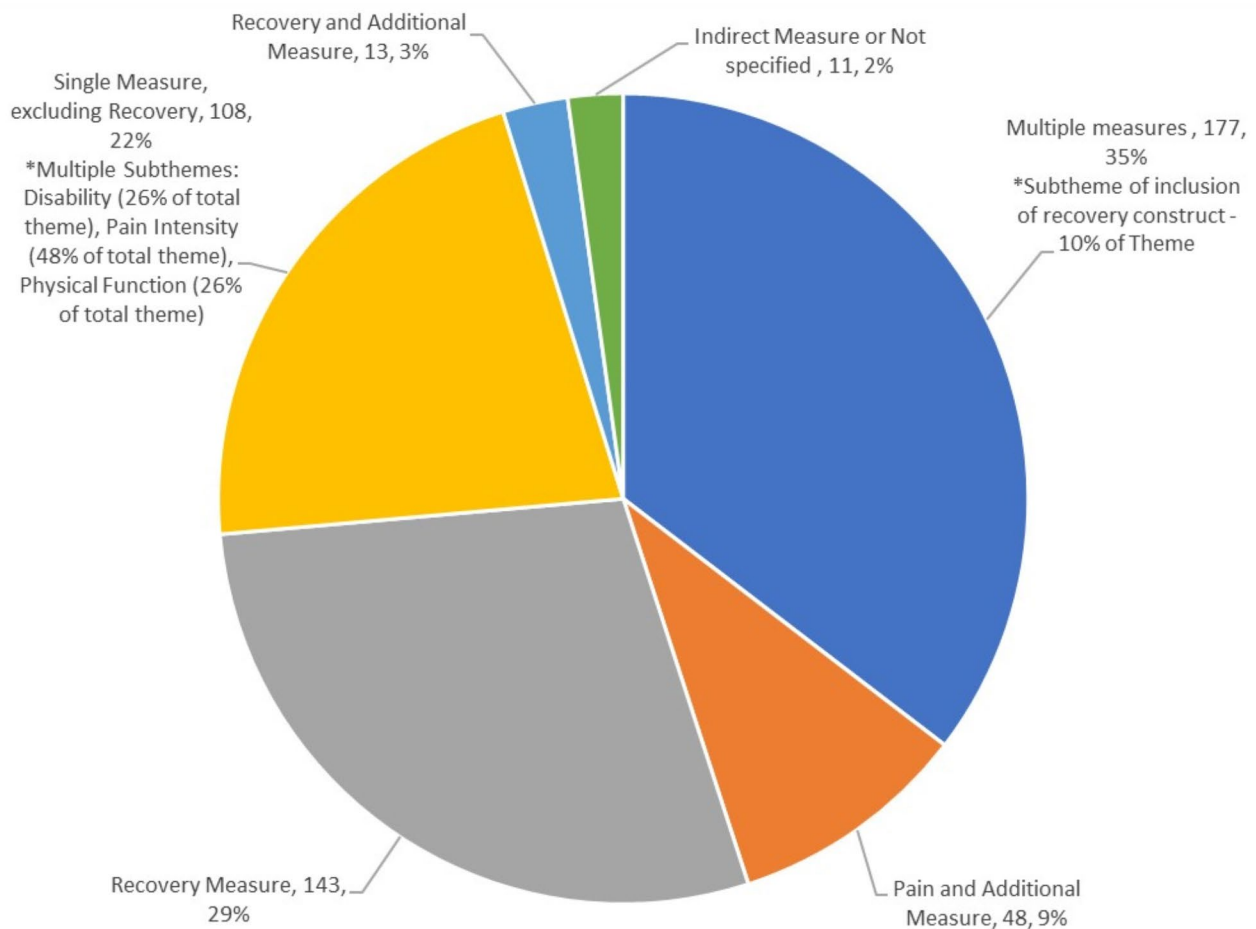


Fig. 3 Six main themes of recovery measurements across entire synthesis. Results are presented in both counts and percentage of total synthesis

Table 2 Themes and sub-themes (in italics) of perspectives of recovery across the entire synthesis

Perspective of recovery	Frequency (n)	%
Patient	149	29.8
<i>Patient - Direct</i>	99	19.8
<i>Patient - Indirect</i>	50	10.0
Provider	333	66.6
<i>Provider Only</i>	136	27.2
<i>Patient but provider analyzed</i>	197	39.4
Patient and Provider	18	3.6

measure, while others included a pain-based measure and a physical function measure.

Indirect measure/Not specified The remainder of the included studies (2%) either used an indirect measure of recovery (e.g., economic evaluation) or were not specified.

Recovery perspectives

Recovery perspectives were analyzed across all studies to determine if the patient and provider were accounted for in recovery definitions and measurements (Table 2).

Three main themes and four sub-themes adequately described the data.

Provider The largest theme was the perspective of the provider, with two sub-themes: Provider only, Patient but provider analyzed. 27.2% (n=136) of studies had included a perspective of recovery from the provider alone. The second sub-theme in this category was when the patient’s perspective was initially included but was further analyzed by the provider in 39.4% of studies. Examples included the use of patient-reported data, but then a threshold was

placed upon to dichotomize patients as recovered or not. Taken together (66.6% of synthesis), the patient's perspective regarding recovery was either not preserved in making a recovery determination or was not made from their perspective.

Patient The second most frequent perspective of recovery gathered was from the patient in 29.8% of studies. Perspectives were either direct (recovery-related measure), or indirect where a patient-reported measure was used but was not recovery-centric (i.e., disability).

Patient and provider Lastly, in 3.6% ($n=17$) of studies, both patient and provider perspectives were gathered. No included studies reported on potential inconsistencies between the perspectives of patients and providers regarding recovery.

See [S3 File](#) for a comprehensive list of all data synthesis for recovery definitions, measurements, and perspectives.

Discussion

This review represents the first attempt to synthesize the measurements and definitions of recovery for LBP across quantitative and qualitative literature, while also seeking to identify how perspectives of recovery may align or differ between patients and providers. Seminal work was completed in this area by Kamper and colleagues over 10 years ago [6], and it appears that the field has progressed very little given the numerous definitions and measures of recovery identified through thematic analysis in the present review. At the time of writing, recovery is not emphasized as a highly relevant domain in LBP research and is not included as part of a core outcome set for LBP [37, 38]. While a lack of standardization with respect to defining and measuring recovery cannot solely be blamed for the current burden of LBP [1, 39], it is very likely a contributing factor. Inconsistent definition of outcomes can lead to under-prioritization and a lack of funding, which compounds challenges with clinical recommendations [20, 39]. As a result, urgent action is required from researchers, clinicians, and patients to come together to help operationalize the concept of recovery.

Recovery definitions

Alarming, this review highlights that more than 1/3 of published studies lacked a formal definition of recovery. This issue is further complicated by almost half of the included studies being low quality indicating that the field does not have a consistent understanding of the construct of recovery. As a result, it becomes challenging to recommend definitions or measurements of recovery based on the methodological quality of included studies. It should also be noted that definition of different yet

related concepts like recurrence also are limited due in part to a lack of agreed upon definitions for recovery in patients with LBP [40].

This review identified **seven main themes in which recovery has been defined**. Many of these themes appear to be highly emblematic of an ableist normal, where the attainment of normal is viewed as the goal, and any failure to do is likely to discriminate against those who cannot achieve it [41]. While issues of ableism are likely to persist in defining recovery, ceasing to use singular cut-off scores could assist in de-coupling a patient's recovery and notions of normality [42]. It should also be noted that Kamper and colleagues also highlighted that dichotomous definitions of recovery were the most frequently adopted method for defining recovery [6], as opposed to defining recovery as a continuous process. In contrast, our work found that 20% of included studies defined recovery compared to a singular-cut off score. As a result, it appears that the field is shifting, but this practice remains consistent 10 years since the last review.

Efforts for recommendations on recovery definition have been made but are not widely implemented, given the multiple themes of definitions noted in the present study. Previous literature identified the notion that recovery can include several domains depending on whether LBP is acute or chronic, leading to an emphasis on readjustment rather than resolution [43]. Walton and colleagues postulated that recovery could be approximated by self-report and may represent the achievement of a health state perceived by an individual to provide the health resources necessary to achieve an actualised version of themselves [11]. In other words, self-report for recovery could represent one of the more valid definitions for recovery as it would be specific on an individual basis. For such an understanding to take place, multiple understandings of what recovery constitutes are required. For example, for some patients with LBP a return to a pre-injury state or the absence of clinical symptoms may be desirable. For others, a change in a score from a baseline or a triangulation of scores may seem apt. We recommend that a consensus-driven definition of recovery should include a range of accepted definitions given that our work supports that there is no singular definition for recovery of LBP.

Finally, another consideration for the concept of recovery (and its definition) is response shift. Response shift can be defined as a change in the meaning of self-evaluation and is particularly evident in chronic conditions [44, 45]. In the context of potentially chronic conditions such as LBP, response shift could represent a coping strategy used by patients to re-evaluate or re-define the goals that are meaning to them in their course of care [11, 45]. While methods to control for response shift are beyond the scope of this article, the phenomenon itself

is worth mentioning in the context of defining recovery for patients with LBP. As others have indicated, response shift can be one reason why rehabilitation efforts can appear minimal in those with chronic presentations, and response shift should be considered especially for studies focused on treatment effects [46].

Recovery measures

Without a unified definition of recovery, there is currently no accepted gold standard for recovery measurement [47], and measurement of recovery remains problematic. This review highlighted that there were **six main themes to indicate how recovery had been measured** in patients with LBP. One quarter of included studies used a non-recovery centric measure (e.g., pain intensity or disability) in isolation for determination of recovery. Pain and Disability-centric measures have been previously reported to be some of the most common measures in recovery [47], which is consistent with this review. In the example of disability, other author groups have highlighted that patients may demonstrate an improved RMDQ score yet view themselves globally worse [48]. Almost 1/3 of included studies in this review employed the use of a recovery-centric measure such as the GRC. While such measures are known to be easy to administer and can offer excellent test-retest reliability [49, 50], the validity of GRC measures are often questioned as well as their susceptibility to recall bias [48, 50].

Likely related to a lack of a gold standard [47], the largest theme captured in this review for recovery measurement pertained to the triangulation of measures to make a recovery determination. If what is emphasized as important for recovery in patients with LBP is different to different people, than multiple measures are likely needed to properly capture it. As has been previously postulated, this approach seems to be reasonable as at least two sources of information provide overlap, while three sources can aid in the confidence of making a recovery determination through triangulation of findings [51]. Since most included studies were of either low/moderate quality, specific measures cannot be endorsed, but a triangulation approach seems appropriate given current limitations. Thus, our work would recommend the use of multiple measures to triangulate findings when measuring recovery.

In practice, this could look very different between providers. Triangulation of recovery measurement could appear different from patient to patient as a clinically important effect is specific to the individual patient [52]. The use of multiple measures also assists in overcoming the limitations of only evaluating one outcome measure, and risk minimizing the patient experience [53]. Selection of outcomes that highlight multiple domains/attributes is one way in which (a) a provider can be more

certain that there is change in presentation and (b) measure domains of the recovery experience that are important to the patient [54]. Fundamentally, evaluation of recovery should incorporate the patient's perspective as many legacy PROMs were not designed with patients in mind [53–55], and the use of multiple measures enables a comprehensive narrative of a patient's recovery journey.

Perspectives of recovery

We did not find any evidence of perspective of recovery between patients and providers differing in the same study, but the perspectives of patients were mostly not preserved in making recovery determinations. While PROMs were the most frequently used clinical measurements related to recovery, the actual determination of recovery was made by investigators usually in relation to a cut-off score, a combination of cut-off scores, the absence of pain or disability, or improvement from a baseline measurement. The results of this review are consistent with previous criticisms of the field, in that patient perspectives on recovery determinations are not traditionally included beyond self-report measures, suggesting additional measures are needed [56]. Both depression studies and post-surgical literature have reported that patient perspectives can be a blind spot in understanding recovery [13, 57, 58]. As current pain theories like the Multimodal Assessment Model of Pain recognize patient narrative as the best approximate proxy for the pain experience [59], then recovery narratives are also likely the best proxy for the recovery experience. Given that patient perspectives were not incorporated in legacy measures [55], establishment of a consensus regarding recovery that states the importance of patient narrative could advance the field like the proposed International Association for the Study of Pain definition for pain [60, 61].

Strengths and limitations

This review has several strengths. This MSSR synthesizes definitions, measurements, and perspectives of recovery across a wide spectrum of research and included a formal quality assessment of included studies. However, this review is not without limitations. There were fewer qualitative and mixed methods studies included in the synthesis in comparison to quantitative literature, highlighting a gap in existing evidence. As we only included studies that specifically mentioned recovery it is possible that studies that used similar related terms (e.g., Resolution, Improvement, etc.) could have been missed. However, as this review included nearly 500 studies, this omission is very unlikely to significantly alter results. Lastly, another limitation to our work pertains to the thresholds used for quality assessment of the MMAT, as there are no established thresholds for this tool.

Conclusion

Within LBP research, recovery as an outcome is not properly defined, is measured using multiple measures and does not typically incorporate patient perspectives. We recommend that a consensus-driven approach, such as the Delphi method [62] is needed to define a range of accepted definitions for recovery given that our work supports that there is no singular definition for recovery of LBP. Based on the results of this review we also recommend using multiple measures that triangulate findings and including the patient's perspective through qualitative methods. Urgent action is needed to inform a range of accepted definitions to accommodate the heterogeneous presentation of LBP.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12891-024-07892-2>.

Supplementary Material 1: Prisma Checklist
Supplementary Material 2: Search Strategy
Supplementary Material 3: Data Synthesis File
Supplementary Material 4
Supplementary Material 5
Supplementary Material 6
Supplementary Material 7
Supplementary Material 8
Supplementary Material 9

Acknowledgements

N/A.

Author contributions

MJL, KK, MS, NP, ABR, and DMW contributed to the design of the review. MJL, NP, and MM were responsible for screening articles, data extraction and quality assessment. MJL, NP, and KJ wrote the first draft of the manuscript. All authors have approved and contributed to the final manuscript. MJL is the guarantor of the review.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not for profit-sectors.

Data availability

Data is provided within the manuscript or supplementary information files.

Declarations

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹School of Physical Therapy, Western University, London, ON N6G 1H1, Canada

²London Health Sciences Centre, Health Disciplines, London, Canada

³Spine Care Center, Wakayama Medical University Kihoku Hospital, 219

Myoji, Katsuragi-cho, Ito-gun, Wakayama 649-7113, Japan

⁴Western Libraries, Western University, London, Canada

⁵Faculty of Science, Western University, London, ON, Canada

Received: 24 June 2024 / Accepted: 23 September 2024

Published online: 01 October 2024

References

1. Chen S, Chen M, Wu X, Lin S, Tao C, Cao H, et al. Global, regional and national burden of low back pain 1990–2019: a systematic analysis of the Global Burden of Disease study 2019. *J Orthop Translat*. 2022;32:49–58. <https://doi.org/10.1016/J.JOT.2021.07.005>.
2. The burden of. Musculoskeletal conditions at the start of the new millennium. *World Health Organ Tech Rep Ser* 2003;919.
3. Perrot S, Allaert FA, Concas V, Laroche F. When will I recover? A national survey on patients' and physicians' expectations concerning the recovery time for acute back pain. *Eur Spine J*. 2009;18:419. <https://doi.org/10.1007/S00586-008-0868-6>.
4. Lukacs MJ, Kowalski KL, Peters N, Stanley M, Rushton AB. How is recovery defined and measured in patients with low back pain? Protocol for a mixed study systematic review. *BMJ Open*. 2022;12:e061475. <https://doi.org/10.1136/BMJOPEN-2022-061475>.
5. Hush JM, Kamper SJ, Stanton TR, Ostelo R, Refshauge KM. Standardized measurement of Recovery from nonspecific back Pain. *Arch Phys Med Rehabil*. 2012;93:849–55. <https://doi.org/10.1016/J.APMR.2011.11.035>.
6. Kamper SJ, Stanton TR, Williams CM, Maher CG, Hush JM. How is recovery from low back pain measured? A systematic review of the literature. *Eur Spine J*. 2011;20:9–18. <https://doi.org/10.1007/s00586-010-1477-8>.
7. Modarresi S, Walton DM. Reliability, discriminative accuracy, and an exploration of response shift as measured using the satisfaction and recovery index over 12 months from musculoskeletal trauma. *Musculoskelet Sci Pract*. 2021;51:102300. <https://doi.org/10.1016/J.JMSKSP.2020.102300>.
8. Walton DM, Marsh J. The Multidimensional Symptom Index: a new patient-reported outcome for pain phenotyping, prognosis and treatment decisions. *Eur J Pain (United Kingdom)*. 2018;22:1351–61. <https://doi.org/10.1002/ejp.1224>.
9. Mehling WE, Gopisetty V, Acree M, Pressman A, Carey T, Goldberg H, et al. Acute low back pain and primary care: how to define recovery and chronicity? *Spine (Phila Pa 1976)*. 2011;36:2316–23. <https://doi.org/10.1097/BRS.0b013e31820c01a6>.
10. Borghuis EMC, Reneman MF, Schiphorst Preuper HR. Assessing discrepancies in outcomes of pain rehabilitation: these questionnaires don't measure results that are relevant to me. *Disabil Rehabil* 2020;42. <https://doi.org/10.1080/09638288.2018.1561956>
11. Walton DM, Macdermid JC, Nielson W. Recovery from acute injury: clinical, methodological and philosophical considerations. *Disabil Rehabil*. 2010;32:864–74. <https://doi.org/10.3109/09638280903349511>.
12. Walton D. A review of the definitions of recovery used in prognostic studies on whiplash using an ICF framework. *Disabil Rehabil*. 2009;31:943–57. <https://doi.org/10.1080/09638280802404128>.
13. Rushton A, Jadhakhan F, Masson A, Athey V, Staal JB, Verra ML, et al. Patient journey following lumbar spinal fusion surgery (FuJourn): a multicentre exploration of the immediate post-operative period using qualitative patient diaries. *PLoS ONE*. 2020;15. <https://doi.org/10.1371/JOURNAL.PONE.0241931>.
14. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol*. 2009;62:1006–12. <https://doi.org/10.1016/j.jclinepi.2009.06.005>.
15. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015: elaboration and explanation. *BMJ (Online)*. 2015;349. <https://doi.org/10.1136/bmj.g7647>.
16. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4:1–9. <https://doi.org/10.1186/2046-4053-4-1>.
17. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The mixed methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Educ Inform*. 2018;34:285–91. <https://doi.org/10.3233/EFI-180221>.

18. Pluye P, Gagnon MP, Griffiths F, Johnson-Lafleur J. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. *Int J Nurs Stud*. 2009;46:529–46. <https://doi.org/10.1016/j.ijnurstu.2009.01.009>.
19. Hong QN, Gonzalez-Reyes A, Pluye P. Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the mixed methods Appraisal Tool (MMAT). *J Eval Clin Pract*. 2018;24:459–67. <https://doi.org/10.1111/JEP.12884/FORMAT/PDF>.
20. Rushton A, Calvert M, Wright C, Freemantle N. Physiotherapy trials for the 21st century - time to raise the bar? *J R Soc Med*. 2011;104. <https://doi.org/10.1258/jrsm.2011.110109>.
21. Yu J, Li X, Li Y, Sun X. Quality of reporting in surgical randomized clinical trials. *Br J Surg*. 2017;104. <https://doi.org/10.1002/bjs.10331>.
22. Pluye P, Hong QN. Combining the Power of Stories and the Power of Numbers: Mixed Methods Research and Mixed Studies Reviews. <http://DxDoiOrg/101146/Annurev-Publhealth-032013-182440>. 2014;35:29–45. <https://doi.org/10.1146/ANNUREV-PUBLHEALTH-032013-182440>
23. Hong QN, Pluye P, Bujold M, Wassef M. Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Syst Rev*. 2017;6:61. <https://doi.org/10.1186/S13643-017-0454-2/TABLES/6>.
24. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol*. 2008;8. <https://doi.org/10.1186/1471-2288-8-45>.
25. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3:77–101. <https://doi.org/10.1191/1478088706qp0630a>.
26. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. *Ann Rheum Dis*. 1978;37:378–81. <https://doi.org/10.1136/ARD.37.4.378>.
27. Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther*. 1996;76. <https://doi.org/10.1093/ptj/76.10.1109>.
28. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)*. 1983;8:141–4. <https://doi.org/10.1097/00007632-198303000-00004>.
29. Goel A, Jadhav N, Shah A, Rai S, Vutha R, Dandpat S, et al. Adult idiopathic de novo lumbar scoliosis: analysis of surgical treatment in 14 patients by only fixation. *J Craniovertebr Junction Spine*. 2020;11. https://doi.org/10.4103/jcvjs.JCVJS_61_20.
30. Altun I, Yüksel KZ. Lumbar herniated disc: spontaneous regression. *Korean J Pain*. 2017;30. <https://doi.org/10.3344/kjp.2017.30.1.44>.
31. da Silva T, Macaskill P, Kongsted A, Mills K, Maher CG, Hancock MJ. Predicting pain recovery in patients with acute low back pain: updating and validation of a clinical prediction model. *Eur J Pain (United Kingdom)*. 2019;23. <https://doi.org/10.1002/ejp.1308>.
32. Hancock MJ, Kjaer P, Kent P, Jensen RK, Jensen TS. Is the number of different MRI findings more strongly Associated with Low Back Pain Than single MRI findings? *Spine (Phila Pa 1976)*. 2017;42. <https://doi.org/10.1097/BRS.0000000000002102>.
33. Angel S, Jensen LD, Maribo T, Gonge BK, Buus N. Narratives of life with long-term low back pain: a follow up interview study. *Scand J Pain* 2017;17. <https://doi.org/10.1016/j.sjpain.2017.09.018>
34. Wilkinson MJB. Does 48 hours' bed rest influence the outcome of acute low back pain? *Br J Gen Pract* 1995;45.
35. Watanabe K, Hosoya T, Shiraiishi T, Matsumoto M, Chiba K, Toyama Y. Lumbar spinous process-splitting laminectomy for lumbar canal stenosis. Technical note. *J Neurosurg Spine*. 2005;3. <https://doi.org/10.3171/spi.2005.3.5.0405>.
36. Fairbank JCT, Davies JB, Couper J, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66. https://doi.org/10.1007/978-3-642-28753-4_101584
37. Chiarotto A, Boers M, Deyo RA, Buchbinder R, Corbin TP, Costa LOP, et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. *Pain*. 2018;159:481. <https://doi.org/10.1097/J.PAIN.0000000000001117>.
38. 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. <https://doi.org/10.1007/S00586-015-3892-3/TABLES/4>.
39. Hoy D, March L, Brooks P, Woolf A, Blyth F, Vos T, et al. Measuring the global burden of low back pain. *Best Pract Res Clin Rheumatol*. 2010;24. <https://doi.org/10.1016/j.berh.2009.11.002>.
40. Stanton TR, Latimer J, Maher CG, Hancock M. Definitions of recurrence of an episode of low back pain: a systematic review. *Spine (Phila Pa 1976)* 2009;34. <https://doi.org/10.1097/BRS.0b013e318198d073>
41. Nixon SA. The coin model of privilege and critical allyship: implications for health. *BMC Public Health*. 2019;19:1–13. <https://doi.org/10.1186/s12889-019-7884-9>.
42. Nixon SA, Yeung E, Shaw JA, Kuper A, Gibson BE. Seven-step Framework for critical analysis and its application in the field of physical therapy. *Phys Ther*. 2017;97:249–57. <https://doi.org/10.2522/ptj.20160149>.
43. Myburgh C, Boyle E, Lauridsen HH, Hestbaek L, Kongsted A. What influences retrospective self-appraised recovery status among danes with low-back problems? A comparative qualitative investigation. *J Rehabil Med*. 2015;47:741–7. <https://doi.org/10.2340/16501977-1987>.
44. Barclay-Goddard R, Epstein JD, Mayo NE. Response shift: a brief overview and proposed research priorities. *Qual Life Res*. 2009;18. <https://doi.org/10.1007/s11136-009-9450-x>.
45. Sprangers MAG, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. *Soc Sci Med*. 1999;48. [https://doi.org/10.1016/S0277-9536\(99\)00045-3](https://doi.org/10.1016/S0277-9536(99)00045-3).
46. Nagl M, Farin E. Response shift in quality of life assessment in patients with chronic back pain and chronic ischaemic heart disease. *Disabil Rehabil*. 2012;34. <https://doi.org/10.3109/09638288.2011.619616>.
47. Kamper SJ, Maher CG, Herbert RD, Hancock MJ, Hush JM, Smeets RJ. How little pain and disability do patients with low back pain have to experience to feel that they have recovered? *Eur Spine J*. 2010;19:1495. <https://doi.org/10.1007/S00586-010-1366-1>.
48. Froud R, Ellard D, Patel S, Eldridge S, Underwood M. Primary outcome measure use in back pain trials may need radical reassessment. *BMC Musculoskelet Disord*. 2015;16:1–11. <https://doi.org/10.1186/S12891-015-0534-1/FIGURES/1>.
49. Kamper SJ, Maher CG, Mackay G. Global rating of Change scales: a review of strengths and weaknesses and considerations for Design. *J Man Manip Ther*. 2009;17:163. <https://doi.org/10.1179/JMT.2009.17.3.163>.
50. Bobos P, Ziebart C, Furtado R, Lu Z, MacDermid JC. Psychometric properties of the global rating of change scales in patients with low back pain, upper and lower extremity disorders. A systematic review with meta-analysis. *J Orthop*. 2020;21:40. <https://doi.org/10.1016/J.JOR.2020.01.047>.
51. Walton DM, Elliott JM. A new clinical model for facilitating the development of pattern recognition skills in clinical pain assessment. *Musculoskelet Sci Pract*. 2018;36:17–24. <https://doi.org/10.1016/j.msksp.2018.03.006>.
52. Abdel Shaheed C, Mathieson S, Wilson R, Furnage AM, Maher CG. Who should judge treatment effects as unimportant? *J Physiother* 2023;69. <https://doi.org/10.1016/j.jphys.2023.04.001>
53. Hickey M, Barry D, Redito J, Anand N, Bianchi N, Reneman M, et al. Content validity of patient-reported outcome measures for patients with chronic pain: considering the patient's perspective. *Pain*. 2023;164. <https://doi.org/10.1097/j.pain.0000000000002714>.
54. Reneman MF, Brandsema KPD, Schrier E, Dijkstra PU, Krabbe PFM. Patients first: toward a patient-centered instrument to measure impact of chronic pain. *Phys Ther*. 2018;98. <https://doi.org/10.1093/ptj/pzy040>.
55. Cook CE, Wright A, Wittstein J, Barbero M, Tousignant-Lafamme Y. Five recommendations to address the limitations of patient-reported outcome measures. *J Orthop Sports Phys Therapy*. 2021;51:562–5. <https://doi.org/10.2519/JOSPT.2021.10836>.
56. Hush JM, Refshaug K, Sullivan G, De Souza L, Maher CG, McAuley JH. Recovery: what does this mean to patients with low back pain? *Arthritis Care Res (Hoboken)*. 2009;61:124–31. <https://doi.org/10.1002/art.24162>.
57. Rajabiyazdi F, Alam R, Pal A, Montanez J, Law S, Pecorelli N, et al. Understanding the meaning of recovery to patients undergoing abdominal surgery. *JAMA Surg*. 2021;156:758–65. <https://doi.org/10.1001/JAMASURG.2021.1557>.
58. van Grieken RA, Kirkenier ACE, Koeter MWJ, Nabitz UW, Schene AH. Patients' perspective on self-management in the recovery from depression. *Health Expect*. 2015;18:1339–48. <https://doi.org/10.1111/HEX.12112>.
59. Wideman TH, Edwards RR, Walton DM, Martel MQ, Hudon A, Seminowicz DA. The Multimodal Assessment Model of Pain: a Novel Framework for further integrating the Subjective Pain Experience within Research and Practice. *Clin J Pain*. 2019;35:212. <https://doi.org/10.1097/AJP.0000000000000670>.
60. Vader K, Bostick GP, Carlesso LC, Hunter J, Mesaroli G, Perreault K, et al. The revised IASP Definition of Pain and Accompanying notes: considerations for the Physiotherapy Profession. *Physiother Can*. 2021;73:103–6. <https://doi.org/10.3138/PTC-2020-0124-GEE>.

61. Raja SN, Carr DB, Cohen M, Finnerup NB, Flor H, Gibson S et al. The revised IASP definition of pain: concepts, challenges, and compromises. *Pain* 2020;161:1976. <https://doi.org/10.1097/J.PAIN.0000000000001939>
62. Skulmoski GJ, Skulmoski GJ, Hartman FT, Krahn J. The Delphi Method for Graduate Research. *J Inform Technol Education: Res.* 2007;6:1–21.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.