Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

Title: Construct Validity of Movement-Evoked Pain Operational Definitions in Older Adults with Chronic Low Back Pain

Running Title: Comparing Movement-Evoked Pain Operational Definitions

Authors: Patrick J. Knox, PT, DPT[1], Corey B. Simon, PT, DPT, PhD[2], Ryan T. Pohlig, PhD [3,4], Jenifer M. Pugliese, PT, DPT[1], Peter C. Coyle, PT, DPT, PhD[1], Jaclyn M. Sions, PT, DPT, PhD[1], Gregory E. Hicks, PT, PhD, FAPTA[1]

[1]Department of Physical Therapy, University of Delaware, Newark, DE. [2]Department of Orthopaedic Surgery, Physical Therapy Division, Duke University, Durham, NC. [3]Department of Epidemiology, University of Delaware, Newark, DE. [4]Biostatistics Core, University of Delaware, Newark, DE.

Corresponding Author: Gregory E. Hicks, PT, PhD, FAPTA; Department of Physical Therapy, University of Delaware; 540 S. College Ave, Newark, DE 19713; Email: ghicks@udel.edu; Phone: 302-831-8910; Fax: 302-831-4234

Conflicts of Interest and Source of Funding: This work was supported by Award Number R01AG0412202 from the National Institute on Aging of the National Institutes of Health. Manuscript preparation was partially supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health (grant number T32-HD007490), as well as a Promotion of Doctoral Studies Scholarship Level I from the Foundation for Physical Therapy Research and the University Doctoral Fellowship Award through the University of Delaware's Graduate College. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Foundation for Physical Therapy Research, or the University of Delaware's Graduate

College. Otherwise, the authors declare that they have no financial relationships, nor other conflicts of interest, to disclose.

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

ABSTRACT

Objective: Movement-evoked pain (MeP) may predispose the geriatric chronic low back pain (LBP) population to health decline. Since there are differing operational definitions for MeP, the question remains as to whether these different definitions have similar associations with health outcomes in older adults with chronic LBP.

Design: Cross-sectional analysis of an observational study.

Setting: Clinical research laboratory.

Subjects: 226 older adults with chronic LBP.

Methods: This secondary analysis used baseline data from a prospective cohort study (n=250). LBP intensity was collected before and after the Repeated Chair Rise Test, Stair Climbing Test, and Six-Minute Walk Test; MeP change scores (i.e., sum of pretest pain subtracted from posttest pain) and aggregated posttest pain (i.e., sum of posttest pain) variables were calculated. LBP-related disability and self-efficacy were measured by the Quebec Back Pain Disability Scale (QBPDS) and Low Back Activity Confidence Scale (LOBACS), respectively. Physical function was measured with the Health ABC Performance Battery. Robust regression with HC3 standard errors was used to evaluate adjusted associations between both MeP variables and disability, self-efficacy and physical function.

Results: Greater aggregated posttest MeP was independently associated with worse disability ($b=0.593$, $t=2.913$, $p=0.004$), self-efficacy ($b=-0.870$, $t=-3.110$, $p=0.002$), and physical function ($b=-0.017$, $t=-2.007$, $p=0.039$); MeP change scores were not associated with any outcome (all $p>0.050$).

Conclusions: Aggregate posttest MeP, but not MeP change scores, was linked to poorer health outcomes in older adults with chronic LBP. Future studies should consider that the construct validity of MeP paradigms partially depends on the chosen operational definition.

Key words: chronic pain, low back pain, movement-evoked pain, pain measurement, geriatrics

INTRODUCTION

Advancing age is a risk factor for high-impact chronic pain (HICP, or pain on most days of the week for ≥3 months that majorly limits at least 1 activity of daily living), which greatly contributes to disability and healthcare costs.(1) The general population is aging,(2) and factors such as advancing age and disability contribute to greater annual healthcare expenditure.(3, 4) Low back pain (LBP) is the most prevalent form of HICP among older adults in the United States,(5, 6) and is associated with markedly decreased function and overall health.(7-11) LBP burden and resultant healthcare expenditure are therefore expected to continue increasing in parallel with the aging population. To improve health outcomes for older adults with HICP, it is critical to understand which aspects of the pain experience may lead to disability.

Attempts to understand the association between pain and disability have been hampered by the manner in which pain is quantified. Compared to young- and middle-aged adults with chronic LBP, typical measures of pain, such as measures of pain at rest (i.e., resting pain) and pain from prior experiences (i.e., recall pain,) tend to be lower and less related to disability among older adults with chronic LBP.(12, 13) Further, resting and recall measures do not distinguish between different aspects of pain, limiting the ability to isolate which component of the pain experience is contributing to disability. As LBP is most restrictive for older adults during functional activities, (14, 15) it is critical to investigate if pain exacerbated by movement (i.e, movement-evoked pain) is associated with health outcomes in older adults with chronic LBP.

Recent evidence established the utility of standardized physical function tests to elicit movement-evoked pain (MeP) in a cohort of older adults with chronic LBP; MeP, but not resting and recall pain, was associated with poorer self-reported and performance based function.(16) While this investigation demonstrated the functional relevance of this paradigm, the question remains as to whether the aggregate scoring method (i.e., summation of multiple pain ratings across the duration of movements) optimally captures the influence of MeP in this population. A recent review established that MeP can be operationally defined (and by default calculated) in other ways, including: the worst pain experienced with movement, or peak; or a quantified change pre- to post-movement, or change score.(17) Given that research to date has not directly compared different MeP methodologies, it is unclear which operational definition best characterizes the impact of MeP on physical function and disability (i.e., construct validity) in older adults with chronic LBP. This investigation sought to address this knowledge gap by determining whether MeP change scores or aggregate posttest MeP ratings are more strongly associated with health outcomes in older adults with chronic LBP. Therefore, the purpose of the investigation was to explore the construct validity of two MeP operational definitions (i.e., summed MeP change scores and aggregate posttest MeP scores) by examining the strength of their associations with LBP-related disability, self-efficacy and physical function in older adults with chronic LBP. Based on previous reports of the considerable pain variability in LBP populations,(18, 19) and our use of several functional tests to elicit MeP, it is possible for pretest pain scores to sequentially increase and remain elevated across tests; change scores may resultantly mischaracterize MeP responses due to underestimation. Thus, we hypothesized that aggregate posttest MeP scores would be more strongly associated with disability, self-efficacy, and physical function than summed MeP change scores.

METHODS

This study is a secondary analysis of baseline data from a prospective cohort study of community-dwelling older adults with chronic LBP from March 2013 to December 2016 in the northern Delaware region.(20) Participants signed a consent form that was approved by the

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

University of Delaware's Institutional Review Board. Participants provided demographic, anthropometric, and self-reported outcome measure data prior to the initiation of functional testing and concomitant MeP assessment.

Participants

Older adults (60-85 years old) with primary complaints of moderately intense chronic low back pain (i.e., intensity >3/10, frequency ≥4 days/week, duration ≥3 months) were recruited to participate. Individuals were excluded if they had LBP symptoms (e.g., unrelenting night pain) suggestive of a possible serious underlying condition, severely limited functional mobility (e.g., needing more than a cane for household ambulation), a progressive neurological disorder (e.g., Parkinson's disease), terminal illness, or cognitive impairment (i.e., a Mini Mental State Examination score <18/30).(21) Ultimately, 250 individuals were enrolled (Figure 1).

Baseline Characteristics

At the baseline session, participants reported their age and sex, and had their body mass index (BMI) measured and calculated. Pain thermometers (range from 0-10, where 0=no pain and 10=worst pain imaginable) were used to measure resting and recall pain.(22, 23) Resting pain was defined as the current level of pain, whereas recall pain was measured by asking participants to rate their "best" and "worst" pain in the 24 hours preceding baseline evaluation. Resting and recall pain were averaged to form a composite pain intensity variable (hereafter referred to as composite resting and recall pain intensity).(24, 25) Participants also reported the duration of their chronic LBP in units of years.

Movement-evoked Pain

The Repeated Chair Rise (RCR), Stair Climbing (SCT), and Six-minute Walk tests (6MWT) were used to elicit pain; methodological details have been previously published.(16) Each test is a reliable and valid measure of physical function,(26-31) and has previously demonstrated the ability to elicit pain in older adults.(16, 32-34) In short, participants were asked to completely stand from a chair five times for the RCR,(35) ascend and descend 12 steps at a self-selected pace for the SCT,(36) and to ambulate as far as possible within 6 minutes for the 6MWT;(37) participants were asked to report their LBP intensity levels (0=no pain, 10=worst pain imaginable) immediately before and after each performance test. MeP change scores were calculated by subtracting pretest pain from posttest pain for each test, and summing the change scores together. Aggregate posttest MeP was calculated by summing the posttest pain levels together.

LBP-related Disability: Quebec Back Pain Disability Scale (QBPDS)

LBP-related disability was collected with the QBPDS,(38) a reliable and valid measure of condition-specific disability in older adults.(39) The 20-item QBPDS asks individuals to rate how their LBP impacts certain activities of daily living (e.g., walking) on a scale of 0 (not difficult at all) to 5 (unable to do). Scores range from 0-100, where higher scores equate to greater LBP-related disability.

LBP-related Self-efficacy: Low Back Activity Confidence Scale (LOBACS)

LBP-related self-efficacy was collected with the LOBACS, which has established reliability and validity as a measure of condition-specific self-efficacy in LBP populations.(40, 41) The LOBACS includes 15 items that ask individuals with LBP to rate their level of confidence in performing certain activities across three subscales: functional self-efficacy (e.g., confidence in

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

carrying and lifting), self-regulatory self-efficacy (e.g., confidence in caring for LBP), and exercise self-efficacy (e.g., confidence in exercising to maintain health). For each item, responses range from 0%-100% (0%=no confidence, 100%=complete confidence). Subscales are averaged to yield a summary score; lower scores reflect decreased self-efficacy.

Physical Function: Health ABC Performance Battery

Physical function was measured with the Health ABC performance battery, which consists of repeated chair stands, standing balance tasks, gait speed, and a narrow walk test for evaluating dynamic balance.(42) For the RCR component, participants were asked to perform 5 chair stands from sitting as quickly as possible; time needed to complete the stands was measured in seconds. Participants completed this task once; both MeP and transitional mobility were gleaned from the singular trial. Standing balance was measured in one trial across three separate postures (semi-tandem, full tandem, and single leg stance), in which participants were tasked with holding each position for up to 30 seconds. For gait speed, individuals were asked to walk at their self-selected pace along a 6-meter walkway; time was measured in seconds. The same 6-meter walkway was used for the narrow walk test. Individuals were asked to walk at their self-selected pace while keeping their feet within a 20cm wide area; time was measured in seconds, and two trials were averaged. Performance on each test was converted to four ratio scores: chair stands per second, total duration of standing balance divided by total possible time of 90 seconds, gait speed in meters per second, and narrow walk test speed in meters per second. These ratios were summed to yield a continuous scale of function from 0-4, where lower scores indicate poorer function.(42)

Statistical Analysis

All analyses were conducted with SPSS software version 28 (IBM Corp., Armonk, NY, USA). Descriptive statistics were conducted for all study variables; means and standard deviations were calculated for continuous variables, while frequencies were observed for categorical variables. In subsequent analyses, listwise deletion was used for missing data. To test the suitability of this approach, means for study variables were compared between participants with and without missing data. These analyses demonstrated that participants with missing data had worse BMIs, composite resting and recall pain intensity, LBP-related disability, LBP-related self-efficacy, and physical function ($p<0.05$). Given that participants with missing data were unique from the rest of the cohort, and MeP data was most commonly missing, data imputation was deemed inappropriate. Pearson correlations were used to test bivariate correlation relationships between independent variables (i.e., MeP change scores and aggregated posttest MeP) and dependent variables (i.e., disability, self-efficacy, and physical function). Associations between MeP variables and each respective dependent variable were initially evaluated by multiple linear regression models that accounted for five covariates that could confound associations with disability, self-efficacy, and/or physical function: age,(43) sex,(44, 45) BMI,(46) composite resting and recall LBP intensity,(47) and LBP duration.(45) For each of the three models, all variables were entered into one block. Variance inflation factors (VIF) were computed to test for multicollinearity. However, these models violated either normality (examined with Shapiro-Wilk and Kolmogorov-Smirnov tests) or homoscedasticity (examined with White's test) assumptions ($p<0.05$). To address these violations, robust regression with HC3 standard errors(48) was used to model the relationship between MeP variables and LBP-related disability (QBPDS), LBP-related self-efficacy (LOBACS), and physical function (Health ABC performance battery). Significance was set at <0.05 for all analyses.

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

RESULTS

Participants' baseline characteristics are detailed in **Table 1**. Our cohort reported an average age of 69.7 years (±6.8 years) and consisted of marginally more females than males (51.2% vs. 48.8%). The majority of participants were Caucasian (85.5%) and retirees (64.7%). On average, participants reported mild-to-moderate composite resting and recall LBP intensity 3.1 (±1.5). Average MeP change scores were 0.7±2.0, with a range of values from -6 to 11. In turn, aggregated posttest MeP was 5.9±5.4 on average, with a range from 0 to 25. Prevalence estimates of MeP, as defined by change scores or aggregated ratings ≥1, were 41% for MeP change scores and 81.3% for aggregate posttest MeP.

Table 2 contains bivariate correlations between MeP calculation methods and LBP-related disability, self-efficacy, and physical function. Aggregated posttest MeP demonstrated larger correlations with disability and self-efficacy (r=0.414, p<0.001 and r=-0.346, p<0.001, respectively) than MeP change scores (r=0.159, p=0.017 and r=-0.004, p=0.955, respectively). Neither aggregated posttest MeP nor MeP change scores were significantly correlated with physical function; however, aggregated posttest MeP had larger correlation and lower p-values than MeP change scores (r=-0.125, p=0.059 vs. r=-0.048, p=0.471). Though the respective MeP calculations were significantly correlated (r=0.311, p<0.001), the magnitude reflects a moderate effect size,(49) suggesting the calculation methods are related yet distinct.

Adjusted associations between MeP calculation methods and LBP-related disability are presented in **Table 3**; multicollinearity was not present among model inputs (VIFs from 1.023 to 1.698). The overall model accounted for 31.1% of variance in LBP-related disability (F=14.054, adjusted or aR^2=0.289, p=<0.001). Aggregated posttest MeP independently contributed to the model, such that a one-unit increase corresponds to greater perceived disability (b=0.593, t=2.913, p=0.004); MeP change scores were not associated with LBP-related disability (p=0.407). Composite resting and recall pain intensity were also associated with greater perceived disability (b=3.044, t=3.611, p<0.001).

Multicollinearity was not present among inputs to model associations between MeP and LBP-related self-efficacy (VIFs from 1.025 to 1.694); results for these associations are presented in **Table 4**. The overall model accounted for 29.3% of variance in LBP-related self-efficacy (F=12.876, aR^2=0.270, p=<0.001). Aggregated posttest MeP was independently associated with poorer self-efficacy after adjusting for covariates (b=-0.870, t=-3.110, p=0.002), while MeP change scores were not associated with LBP-related self-efficacy (p=0.056). Composite resting and recall pain intensity were also associated with poorer self-efficacy (b=-2.447, t=-2.046, p=0.042).

Table 5 contains adjusted associations between MeP and physical function. Based on VIF values from 1.023 to 1.688, multicollinearity was not present. The overall model accounted for 35.7% of variance in physical function (F=17.424, aR^2=0.336, p=<0.001). Similar to the results for self-reported outcomes, aggregated posttest MeP was associated with poorer physical function (b=-0.017, t=-2.007, p=0.039), while MeP change scores were not associated with physical function (p=0.673). Of note, composite resting and recall pain intensity also were not associated with physical function (p=0.309).

We performed an exploratory post-hoc multivariate analysis of covariance (MANCOVA) to determine the extent to which aggregate versus change scores were associated with a global measure of physical function beyond covariates (i.e., age, sex, BMI, composite resting and recall LBP intensity, and LBP duration), which takes into account complementary but distinct aspects of both self-report and physical performance measures.(50) The residual correlation between

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

QBPDS and Health ABC performance battery scores was modest (-0.283) and thus justifies the rationale for running this additional model. Results from the MANCOVA ($p<0.05$) corroborate findings from individual robust regression models for LBP-related disability and physical function; aggregate posttest MeP explained 4.3% of variance in the global physical function measure after controlling for covariates, while MeP change scores explained 0.6% of variance in the combined dependent variable.

DISCUSSION

This investigation is the first to examine the methodological question of whether distinct operational definitions of MeP differentially influence associations with health outcomes in a high-impact chronic pain population (as defined by Pitcher et al.(1)). Aggregated posttest MeP was independently associated with worse cross-sectional measures of disability, self-efficacy and performance-based physical function in older adults with chronic LBP. Conversely, MeP change scores were not associated with any of the dependent variables. Therefore, our hypotheses were fully supported, demonstrating that different calculation methods may influence the observed associations between MeP and health outcomes. As such, it appears that the specificity and relevance of MeP paradigms in chronic pain populations partially depends on how MeP is operationally defined and subsequently calculated.

Our data demonstrates that aggregate posttest MeP may be a more meaningful method to quantify MeP in older adults with chronic LBP, as it is associated with worse outcomes across distinct domains: LBP-related disability, LBP-related self-efficacy, and a physical performance battery that includes transitional mobility, gait speed, and both static and dynamic balance tasks. This evidence augments prior findings in older adults with chronic LBP, which established a link between aggregate posttest MeP and poorer function as measured by a generalized self-report outcome (i.e., the Late Life Function and Disability Instrument, or LLFDI) as well as a performance-based task (i.e., the Timed Up-and-Go test).(16) Given that worse disability, self-efficacy, and physical function are predictive of poor longitudinal health outcomes,(51-53) aggregate posttest MeP may be a risk factor for health decline in older adults with chronic LBP. Collectively, it appears that aggregate posttest MeP is clinically relevant as evidenced by its associations with a cadre of diverse yet interrelated health outcomes. Therefore, our results support the construct validity of aggregated posttest MeP among older adults with chronic LBP. Future research is warranted to build upon this clinical relevance by investigating if aggregated posttest MeP is associated with other meaningful health outcomes (e.g., health-related quality of life and physical activity levels) in this patient population.

The relative superiority of aggregated posttest MeP as compared to MeP change scores may be due to characteristics of the patient population. Since the factors underlying chronic LBP in older adults are highly heterogeneous,(54, 55) it is unlikely that the same movement will elicit pain across individuals in this population. Therefore, multimodal paradigms (e.g., ambulation and transitional mobility in combination) may induce MeP more reliably than unimodal tasks (e.g., ambulation or transitional mobility in isolation) in this population. The measurement of MeP across several physical function tasks is complicated by the fact that pain behavior is known to be variable in geriatric and LBP populations,(18, 19, 56) and MeP may be driven by peripheral and central pain mechanisms.(17) Due to these factors, pain may fluctuate between tests that are used to elicit MeP, and pain may not linearly or consistently increase during one or all tasks of an MeP paradigm. As such, pretest pain levels may increase during the testing session and result in minimal change in pain with testing done later in the session, thereby yielding resultantly low MeP change scores, even if posttest pain is comparatively greater than resting and recall

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

measures of pain. Thus, summing change scores across tests may misrepresent the occurrence of MeP. Instead, aggregate posttest MeP may better measure the burden associated with MeP among older adults with chronic LBP; replication studies are required to support this theory.

Though the intent of this study was to establish construct validity of distinct operational definitions for MeP, it is important to note that aggregate posttest MeP was the lone variable associated with performance-based physical function. Our results are reinforced by other evidence in older adults with chronic LBP that established aggregate posttest MeP is more strongly related to performance-based measures of physical function than resting and recall pain intensity.(16) These quantitative findings are consistent, and align with other qualitative studies which indicate that LBP is most restricting during activities or movements in the older adult population.(14, 15) Regarding perception based outcomes, a prior study established that aggregate posttest MeP, but not composite resting and recall, was associated with a general measure of perceived function (i.e., LLFDI).(16) In the current study, greater composite resting and recall pain intensity was associated worse QBPDS scores; therefore, resting and recall pain may be more strongly associated with condition-specific (e.g., QBPDS) rather than general measures of perceived function (e.g., LLFDI). Altogether, clinicians and researchers may want to consider capturing both MeP and composite resting and recall pain in their assessments of older adults in order to better characterize how LBP impacts function and disability.

This study has several strengths that should be considered. Most importantly, this study directly addresses a methodological knowledge gap that pertains to all MeP investigations, irrespective of patient population. We expanded evidence for an MeP paradigm that was initially established in older adults with chronic LBP: the practice of using standardized tasks to elicit MeP, combined with the aggregated posttest calculation method, may advance our understanding of pain-related health decline in this population. Relative to other MeP literature, our study included a large cohort of older adults with high-impact chronic pain and incorporated analyses that adequately accounted for factors that may confound observed associations with disability, self-efficacy and physical function. Finally, consistency was observed across the robust regression models, such that aggregated posttest MeP was associated with poorer disability, self-efficacy, and physical function outcomes.

Study results should be viewed in the context of limitations. The study design was cross-sectional and observational; thus, the associations do not amount to causality. Participants with missing data had generally poorer measures of function and disability, suggesting that associations in the study may be biased towards underestimation. Further, our results may be dependent upon the tests used to elicit MeP, as well as the chosen outcomes. It is possible that change scores are more influential in different MeP paradigms, with different health outcomes, or in different patient populations. Additionally, repeated chair rises were utilized to elicit MeP as well as to measure transitional mobility performance. It is possible for pain exacerbation to have interacted with transitional mobility performance; however, this interaction is actually a strength of the study, as it provides context for how MeP may impart disability in this population. Further, the other components of the Health ABC performance battery were not intended to elicit MeP; thus, the finding that greater MeP is associated with poorer physical function cannot be entirely driven by repeated chair rises. Finally, our cohort had mild-to-moderate pain characteristics and were recruited for study participation rather than seeking medical care, which may impact the generalizability of findings.

Conclusion

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

In older adults with chronic LBP, the amount of pain reported following multimodal functional testing may better represent MeP than the amount that pain changes as a result of functional testing. These findings have both clinical and research implications; clinicians should view MeP as a characteristic that is unique from resting and recall pain, while researchers should consider how the validity of MeP paradigms is related to the procedures that evoke pain as well as the operational definition (i.e., the calculation method) of MeP. Future research is needed to corroborate if aggregate posttest MeP is associated with poorer health outcomes in other cohorts of older adults with chronic LBP, and to determine whether aggregate posttest MeP is detrimental in other geriatric chronic pain populations (e.g., lower extremity osteoarthritis).

REFERENCES

1. Pitcher MH, Von Korff M, Bushnell MC, Porter L. Prevalence and Profile of High-Impact Chronic Pain in the United States. *J Pain* 2019;**20**(2)**:** 146-60.

2. He W, Goodkind D, Kowal PR. An aging world: 2015. United States Census Bureau Washington, DC, 2016.

3. Hayden JA, Chou R, Hogg-Johnson S, Bombardier C. Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol* 2009;**62**(8)**:** 781-96.e1.

4. Gaskin DJ, Richard P. The economic costs of pain in the United States. *J Pain* 2012;**13**(8)**:** 715-24.

5. Institute of Medicine Committee on Advancing Pain Research C, Education. The National Academies Collection: Reports funded by National Institutes of Health. Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research. Washington (DC): National Academies Press (US)

Copyright © 2011, National Academy of Sciences.; 2011.

6. Freburger JK, Holmes GM, Agans RP, Jackman AM, Darter JD, Wallace AS, Castel LD, Kalsbeek WD, Carey TS. The rising prevalence of chronic low back pain. *Arch Intern Med* 2009;**169**(3)**:** 251-8.

7. Reid MC, Williams CS, Gill TM. Back pain and decline in lower extremity physical function among community-dwelling older persons. *J Gerontol A Biol Sci Med Sci* 2005;**60**(6)**:** 793-7.

8. Rudy TE, Weiner DK, Lieber SJ, Slaboda J, Boston JR. The impact of chronic low back pain on older adults: a comparative study of patients and controls. *Pain* 2007;**131**(3)**:** 293-301.

9. Edmond SL, Felson DT. Function and back symptoms in older adults. *J Am Geriatr Soc* 2003;**51**(12)**:** 1702-9.

10. Makris UE, Fraenkel L, Han L, Leo-Summers L, Gill TM. Restricting back pain and subsequent mobility disability in community-living older persons. *J Am Geriatr Soc* 2014;**62**(11)**:** 2142-7.

11. Hicks GE, Simonsick EM, Harris TB, Newman AB, Weiner DK, Nevitt MA, Tylavsky FA. Trunk muscle composition as a predictor of reduced functional capacity in the health, aging and body composition study: the moderating role of back pain. *J Gerontol A Biol Sci Med Sci* 2005;**60**(11)**:** 1420-4.

12. Manogharan S, Kongsted A, Ferreira ML, Hancock MJ. Do older adults with chronic low back pain differ from younger adults in regards to baseline characteristics and prognosis? *Eur J Pain* 2017;**21**(5)**:** 866-73.

13. Houde F, Cabana F, Léonard G. Does Age Affect the Relationship Between Pain and Disability? A Descriptive Study in Individuals Suffering From Chronic Low Back Pain. *J Geriatr Phys Ther* 2016;**39**(3)**:** 140-5.

14. Makris UE, Higashi RT, Marks EG, Fraenkel L, Gill TM, Friedly JL, Reid MC. Physical, Emotional, and Social Impacts of Restricting Back Pain in Older Adults: A Qualitative Study. *Pain Med* 2017;**18**(7)**:** 1225-35.

15. Edwards TC, Lavallee DC, Bauer Z, Comstock BA, Jarvik JG, Patrick DL, Makris UE, Friedly JL. Problem areas identified as important to older adults with lumbar spinal stenosis. *Spine J* 2015;**15**(7)**:** 1636-44.

16. Knox PJ, Simon CB, Pohlig RT, Pugliese JM, Coyle PC, Sions JM, Hicks GE. A Standardized Assessment of Movement-evoked Pain Ratings Is Associated With Functional Outcomes in Older Adults With Chronic Low Back Pain. *The Clinical Journal of Pain* 2022;**38**(4)**:** 241-49.

17. Fullwood D, Means S, Merriwether EN, Chimenti RL, Ahluwalia S, Booker SQ. Toward Understanding Movement-evoked Pain (MEP) and its Measurement: A Scoping Review. *Clin J Pain* 2021;**37**(1)**:** 61-78.

18. Wesolowicz DM, Bishop MD, Robinson ME. An Examination of Day-to-Day and Intraindividual Pain Variability in Low Back Pain. *Pain Medicine* 2021;**22**(10)**:** 2263-75.

19. Pagé MG, Gauvin L, Sylvestre M-P, Nitulescu R, Dyachenko A, Choinière M. An Ecological Momentary Assessment Study of Pain Intensity Variability: Ascertaining Extent, Predictors, and Associations With Quality of Life, Interference and Health Care Utilization Among Individuals Living With Chronic Low Back Pain. *The Journal of Pain* 2022.

20. Hicks GE, Pohlig RT, Coyle PC, Sions JM, Weiner DK, Pugliese JM, Velasco TO, O'Brien VA. Classification of Geriatric Low Back Pain Based on Hip Characteristics with a 12-Month Longitudinal Exploration of Clinical Outcomes: Findings from Delaware Spine Studies. *Phys Ther* 2021.

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

21. Wood RY, Giuliano KK, Bignell CU, Pritham WW. Assessing cognitive ability in research: use of MMSE with minority populations and elderly adults with low education levels. *J Gerontol Nurs* 2006;**32**(4)**:** 45-54.

22. Herr K, Spratt KF, Garand L, Li L. Evaluation of the Iowa pain thermometer and other selected pain intensity scales in younger and older adult cohorts using controlled clinical pain: a preliminary study. *Pain Med* 2007;**8**(7)**:** 585-600.

23. Ware LJ, Herr KA, Booker SS, Dotson K, Key J, Poindexter N, Pyles G, Siler B, Packard A. Psychometric Evaluation of the Revised Iowa Pain Thermometer (IPT-R) in a Sample of Diverse Cognitively Intact and Impaired Older Adults: A Pilot Study. *Pain Manag Nurs* 2015;**16**(4)**:** 475-82.

24. Jensen MP, Tomé-Pires C, Solé E, Racine M, Castarlenas E, de la Vega R, Miró J. Assessment of pain intensity in clinical trials: individual ratings vs composite scores. *Pain Med* 2015;**16**(1)**:** 141-8.

25. Jensen MP, Castarlenas E, Tomé-Pires C, de la Vega R, Sánchez-Rodríguez E, Miró J. The Number of Ratings Needed for Valid Pain Assessment in Clinical Trials: Replication and Extension. *Pain Med* 2015;**16**(9)**:** 1764-72.

26. Simmonds MJ, Olson SL, Jones S, Hussein T, Lee CE, Novy D, Radwan H. Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. *Spine (Phila Pa 1976)* 1998;**23**(22)**:** 2412-21.

27. Novy DM, Simmonds MJ, Lee CE. Physical performance tasks: what are the underlying constructs? *Arch Phys Med Rehabil* 2002;**83**(1)**:** 44-7.

28. Bohannon RW, Shove ME, Barreca SR, Masters LM, Sigouin CS. Five-repetition sit-to-stand test performance by community-dwelling adults: A preliminary investigation of

times, determinants, and relationship with self-reported physical performance. *Isokinetics and Exercise Science* 2007;**15:** 77-81.

29. Lord SR, Murray SM, Chapman K, Munro B, Tiedemann A. Sit-to-stand performance depends on sensation, speed, balance, and psychological status in addition to strength in older people. *J Gerontol A Biol Sci Med Sci* 2002;**57**(8)**:** M539-43.

30. Lin YC, Davey RC, Cochrane T. Tests for physical function of the elderly with knee and hip osteoarthritis. *Scandinavian journal of medicine & science in sports* 2001;**11**(5)**:** 280-86.

31. Harada ND, Chiu V, Stewart AL. Mobility-related function in older adults: assessment with a 6-minute walk test. *Arch Phys Med Rehabil* 1999;**80**(7)**:** 837-41.

32. Booker S, Cardoso J, Cruz-Almeida Y, Sibille KT, Terry EL, Powell-Roach KL, Riley JL, 3rd, Goodin BR, Bartley EJ, Addison AS, Staud R, Redden D, Bradley L, Fillingim RB. Movement-evoked pain, physical function, and perceived stress: An observational study of ethnic/racial differences in aging non-Hispanic Blacks and non-Hispanic Whites with knee osteoarthritis. *Exp Gerontol* 2019;**124:** 110622.

33. Harden RN, Wallach G, Gagnon CM, Zereshki A, Mukai A, Saracoglu M, Kuroda MM, Graciosa JR, Bruehl S. The osteoarthritis knee model: psychophysical characteristics and putative outcomes. *J Pain* 2013;**14**(3)**:** 281-9.

34. Wideman TH, Finan PH, Edwards RR, Quartana PJ, Buenaver LF, Haythornthwaite JA, Smith MT. Increased sensitivity to physical activity among individuals with knee osteoarthritis: relation to pain outcomes, psychological factors, and responses to quantitative sensory testing. *Pain* 2014;**155**(4)**:** 703-11.

35. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, Scherr PA, Wallace RB. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol* 1994;**49**(2)**:** M85-94.

36. Butler AA, Menant JC, Tiedemann AC, Lord SR. Age and gender differences in seven tests of functional mobility. *J Neuroeng Rehabil* 2009;**6:** 31.

37. Rikli RE, Jones CJ. The Reliability and Validity of a 6-Minute Walk Test as a Measure of Physical Endurance in Older Adults. *Journal of Aging and Physical Activity* 1998;**6**(4)**:** 363-75.

38. Kopec JA, Esdaile JM, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping DL, Williams JI. The Quebec Back Pain Disability Scale. Measurement properties. *Spine (Phila Pa 1976)* 1995;**20**(3)**:** 341-52.

39. Hicks GE, Manal TJ. Psychometric properties of commonly used low back disability questionnaires: are they useful for older adults with low back pain? *Pain Med* 2009;**10**(1)**:** 85-94.

40. Davenport TE, Cleland JA, Yamada KA, Kulig K. Measurement Properties of the Low Back Activity Confidence Scale (LoBACS). *Evaluation & the health professions* 2016;**39**(2)**:** 204-14.

41. Yamada KA, Lewthwaite R, Popovich Jr JM, Beneck GJ, Kulig K, Network PTCR. The Low Back Activity Confidence Scale (LoBACS): preliminary validity and reliability. *Physical therapy* 2011;**91**(11)**:** 1592-603.

42. Simonsick EM, Newman AB, Nevitt MC, Kritchevsky SB, Ferrucci L, Guralnik JM, Harris T. Measuring higher level physical function in well-functioning older adults:

expanding familiar approaches in the Health ABC study. *J Gerontol A Biol Sci Med Sci* 2001;**56**(10)**:** M644-9.

43. Wettstein M, Eich W, Bieber C, Tesarz J. Pain Intensity, Disability, and Quality of Life in Patients with Chronic Low Back Pain: Does Age Matter? *Pain Med* 2019;**20**(3)**:** 464-75.

44. Straight CR, Brady AO, Evans E. Sex-specific relationships of physical activity, body composition, and muscle quality with lower-extremity physical function in older men and women. *Menopause* 2015;**22**(3)**:** 297-303.

45. Rundell SD, Sherman KJ, Heagerty PJ, Mock CN, Dettori NJ, Comstock BA, Avins AL, Nedeljkovic SS, Nerenz DR, Jarvik JG. Predictors of Persistent Disability and Back Pain in Older Adults with a New Episode of Care for Back Pain. *Pain Med* 2017;**18**(6)**:** 1049-62.

46. Hardy R, Cooper R, Aihie Sayer A, Ben-Shlomo Y, Cooper C, Deary IJ, Demakakos P, Gallacher J, Martin RM, McNeill G, Starr JM, Steptoe A, Syddall H, Kuh D. Body mass index, muscle strength and physical performance in older adults from eight cohort studies: the HALCyon programme. *PLoS One* 2013;**8**(2)**:** e56483.

47. Patel KV, Guralnik JM, Dansie EJ, Turk DC. Prevalence and impact of pain among older adults in the United States: findings from the 2011 National Health and Aging Trends Study. *Pain* 2013;**154**(12)**:** 2649-57.

48. Long JS, Ervin LH. Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician* 2000;**54**(3)**:** 217-24.

49. Brydges CR. Effect Size Guidelines, Sample Size Calculations, and Statistical Power in Gerontology. *Innov Aging* 2019;**3**(4).

50. Bean JF, Olveczky DD, Kiely DK, LaRose SI, Jette AM. Performance-based versus patient-reported physical function: what are the underlying predictors? *Phys Ther* 2011;**91**(12)**:** 1804-11.

51. Mutubuki EN, Beljon Y, Maas ET, Huygen FJPM, Ostelo RWJG, van Tulder MW, van Dongen JM. The longitudinal relationships between pain severity and disability versus health-related quality of life and costs among chronic low back pain patients. *Quality of Life Research* 2020;**29**(1)**:** 275-87.

52. Foster NE, Thomas E, Bishop A, Dunn KM, Main CJ. Distinctiveness of psychological obstacles to recovery in low back pain patients in primary care. *PAIN*® 2010;**148**(3)**:** 398-406.

53. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med* 1995;**332**(9)**:** 556-61.

54. Weiner DK, Gentili A, Coffey-Vega K, Morone N, Rossi M, Perera S. Biopsychosocial Profiles and Functional Correlates in Older Adults with Chronic Low Back Pain: A Preliminary Study. *Pain Med* 2019;**20**(7)**:** 1300-10.

55. Weiner DK, Sakamoto S, Perera S, Breuer P. Chronic low back pain in older adults: prevalence, reliability, and validity of physical examination findings. *Journal of the American Geriatrics Society* 2006;**54**(1)**:** 11-20.

56. Dobscha SK, Morasco BJ, Kovas AE, Peters DM, Hart K, McFarland BH. Short-term variability in outpatient pain intensity scores in a national sample of older veterans with chronic pain. *Pain Med* 2015;**16**(5)**:** 855-65.

Figure 1: Consolidated Standards of Reporting Trials (CONSORT) study flow diagram

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

Figure 1

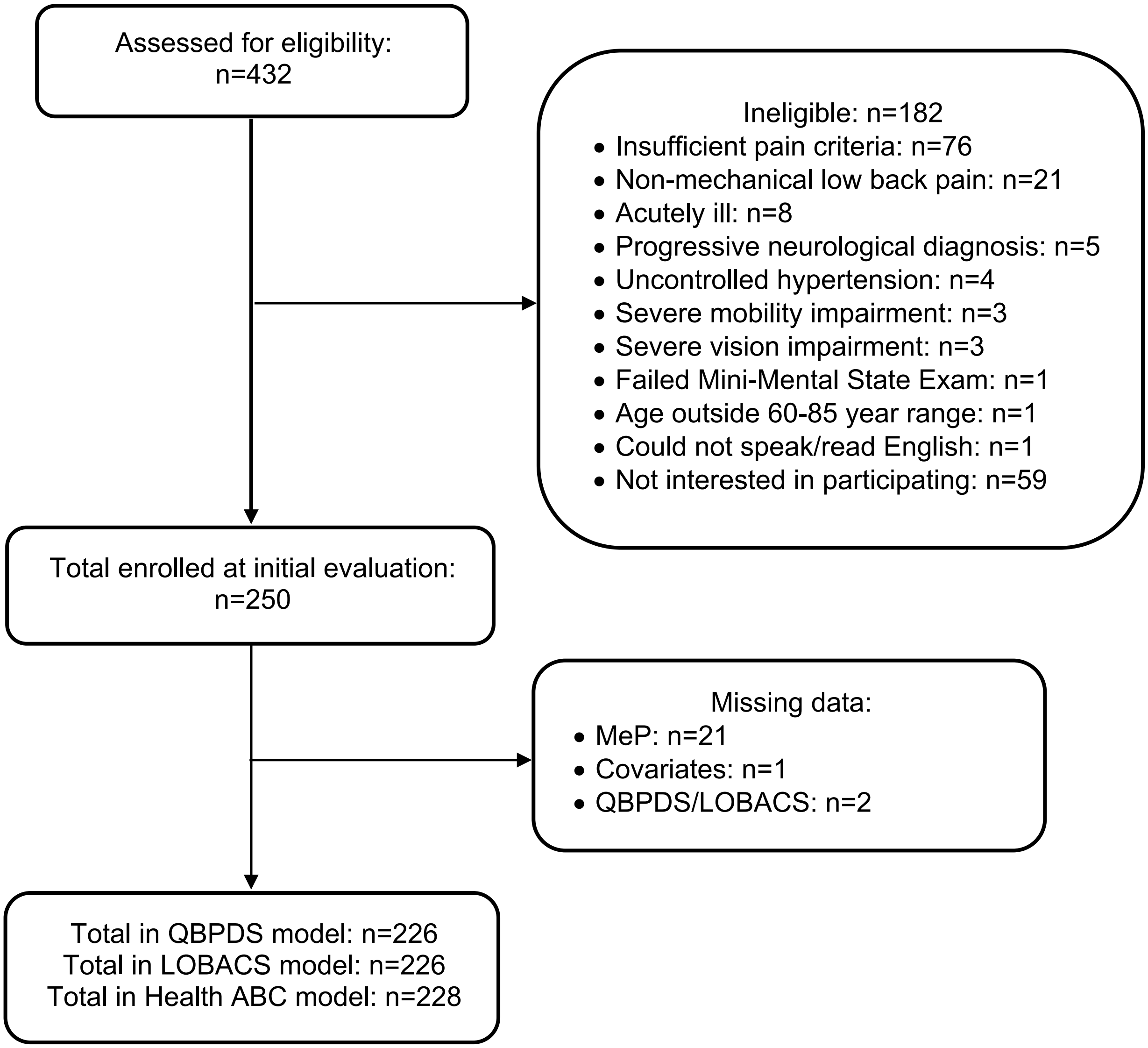

Table 1: Baseline participant characteristics

Variable	n	Mean±SD or n(%)
Age (y)	250	69.7±6.8
Sex (female)	250	128 (51.2%)
Race (Caucasian)	249	213 (85.5%)
Education (graduated college or more)	250	146 (58.4%)
Occupation Status (retired)	249	161 (64.7%)
BMI (kg/m^2)	248	29.4±5.7
LBP onset (years)	250	5.7±7.7
Composite LBP Intensity (0-10)	250	3.1±1.5
MeP Change Scores	229	0.7±2.0
Aggregate Posttest MeP	230	5.9±5.4
QBPDS (0-100)	248	28.3±16.6
LOBACS (0%-100%)	248	65.6±20.4
Health ABC (0-4)	250	2.0±0.6

Abbreviations: BMI = body mass index; LBP = low back pain; MeP = movement-evoked pain; QBPDS = Quebec Back Pain Disability Scale; LOBACS = Low Back Activity Confidence Scale; Health ABC = Health ABC Performance Battery

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

Table 2: Pearson correlations (r) between movement-evoked pain calculation methods and disability, self-efficacy, and physical function.

	MeP Change Score	Aggregate Posttest MeP	QBPDS	LOBACS	Health ABC
MeP Change Score	1				
Aggregate Posttest MeP	0.311**	1			
QBPDS	0.159*	0.414**	1		
LOBACS	-0.004	-0.346**	-0.643**	1	
Health ABC	-0.048	-0.125	-0.392**	0.431**	1

Abbreviations: MeP = movement-evoked pain; QBPDS = Quebec Back Pain Disability Scale; LOBACS = Low Back Activity Confidence Scale; Health ABC = Health ABC Performance Battery

*=p<0.050; **=p<0.010

Table 3: Robust regression output for associations between movement-evoked pain and the Quebec Back Pain Disability Scale

Variable	b	HC3 Standard Error	t-statistic	p-value
Age	0.027	0.143	0.189	0.851
Sex	-7.256	1.896	-3.828	<0.001*
BMI	0.552	0.161	3.428	<0.001*
LBP Duration	-0.098	0.103	-0.953	0.341
Composite LBP Intensity	3.044	0.843	3.611	<0.001*
MeP Change Scores	0.438	0.527	0.831	0.407
Aggregate Posttest MeP	0.593	0.204	2.913	0.004*

Abbreviations: BMI = body mass index; LBP = low back pain; MeP = movement-evoked pain

*=p<0.010; b = unstandardized coefficient; male data displayed for “sex” variable

Pain Medicine
Accepted Manuscript
Version of Record at: https://doi.org/10.1093/pm/pnad034

Table 4: Robust regression output for associations between movement-evoked pain and the Low Back Activity Confidence Scale

Variable	b	HC3 Standard Error	t-statistic	p-value
Age	-0.337	0.171	-1.967	0.050
Sex	11.873	2.457	4.833	<0.001**
BMI	-0.843	0.225	-3.741	<0.001**
LBP Duration	0.105	0.147	0.711	0.478
Composite LBP Intensity	-2.447	1.196	-2.046	0.042*
MeP Change Scores	1.289	0.672	1.919	0.056
Aggregate Posttest MeP	-0.870	0.280	-3.110	0.002**

Abbreviations: BMI = body mass index; LBP = low back pain; MeP = movement-evoked pain

*=p<0.050; **=p<0.010; b = unstandardized coefficient; male data displayed for “sex” variable

Table 5: Robust regression output for associations between movement-evoked pain and the Health ABC Performance Battery

Variable	b	HC3 Standard Error	t-statistic	p-value
Age	-0.041	0.005	-7.950	<0.001**
Sex	0.182	0.065	2.823	0.005**
BMI	-0.31	0.007	-4.525	<0.001**
LBP Duration	0.005	0.005	0.960	0.338
Composite LBP Intensity	0.029	0.029	1.020	0.309
MeP Change Scores	0.007	0.018	0.423	0.673
Aggregate Posttest MeP	-0.017	0.008	-2.077	0.039*

Abbreviations: BMI = body mass index; LBP = low back pain; MeP = Movement-Evoked Pain

*=p<0.050; **=p<0.010; b = unstandardized coefficient; male data displayed for “sex” variable