

TITLE: Classification of Geriatric Low Back Pain Based on Hip Characteristics With a 12-Month Longitudinal Exploration of Clinical Outcomes: Findings From Delaware Spine Studies

RUNNING HEAD: Hip Classifications in Geriatric Low Back Pain

TOC CATEGORY: Geriatrics

ARTICLE TYPE: Original Research

SUBMITTED: January 12, 2021

REVISED: June 3, 2021

ACCEPTED: August 30, 2021

KEYWORDS: Aging, Arthritis, Chronic Pain, Geriatrics, Hip, Hip Osteoarthritis, Low Back Pain, Mobility Limitation, Models, Statistical, Back Pain

AUTHORS: Gregory E. Hicks, PT, PhD, FAPTA¹; Ryan T. Pohlig, PhD^{2,3}; Peter C. Coyle, PT, DPT, PhD¹; J. Megan Sions, PT, DPT, PhD¹; Debra K. Weiner, MD⁴⁻⁸; Jenifer M. Pugliese, PT, DPT¹; Teonette O. Velasco, PT, DPT⁹; Victoria A. O'Brien, MD, PhD¹⁰

¹Department of Physical Therapy, University of Delaware, Newark, DE, ²Department of Epidemiology, University of Delaware, Newark, DE, ³Biostatistics Core, University of Delaware, Newark, DE, ⁴Clinical and Translational Sciences Institute, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, ⁵Geriatric Research Education and Clinical Center, Veterans Affairs Pittsburgh Healthcare System, Pittsburgh, Pennsylvania, ⁶Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, ⁷Department of Anesthesiology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, ⁸Division of Geriatric Medicine, Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, ⁹The Geneva Foundation, National Capital Region, Bethesda, MD, ¹⁰ChristianaCare Health System, Newark, DE.

CORRESPONDING AUTHOR: Gregory E. Hicks, PT, PhD, FAPTA; Department of Physical Therapy, University of Delaware; 540 S. College Ave, Suite 210E, Newark, DE 19713; Email: ghicks@udel.edu ORCID iD: 0000-0001-8611-994X

[H1]ABSTRACT

Objective. The purpose of this study is to identify geriatric chronic LBP subgroups based on the presence of potentially modifiable hip impairments, using Latent Variable Mixture Modeling (LVMM), and to examine the prospective relationship between these subgroups and key outcomes over time.

Methods. Baseline, 3-month, 6-month, and 12-month data were collected from a prospective cohort of 250 community-dwelling older adults with chronic LBP. Comprehensive hip (symptoms, strength, range of motion, and flexibility), LBP (intensity and disability), and mobility function (gait speed and Six-Minute Walk Test) examinations were performed at each timepoint. Baseline hip measures were included in LVMM; observed classes/subgroups were compared longitudinally on LBP and mobility function outcomes using mixed models.

Results. Regarding LVMM, a model with 3 classes/subgroups fit best. Broadly speaking, subgroups were differentiated best by hip strength and symptom presence: subgroup 1 = strong and nonsymptomatic (SNS), subgroup 2 = weak and nonsymptomatic (WNS), subgroup 3 = weak and symptomatic (WS). Regarding longitudinal mixed models, all subgroups improved in most outcomes over time. Specifically, over 12 months, the nonsymptomatic subgroups had lower LBP intensity and disability levels compared to the WS subgroup, whereas the SNS subgroup had better mobility function than the 2 “weak” subgroups.

Conclusion. These subgroup classifications may help in tailoring specific interventions in future trials. Special attention should be given to the WS subgroup, given their consistently poor LBP and mobility function outcomes.

Impact. Among older adults with chronic low back pain, there are 3 hip subgroups: “strong and nonsymptomatic,” “weak and nonsymptomatic,” and “weak and symptomatic.” People in these subgroups demonstrate different outcomes and require different treatment; proper identification will result in tailored interventions designed to benefit individual patients. In particular, people in the WS subgroup deserve special attention, as their outcomes are consistently poorer than those in the other subgroups.

[H1]INTRODUCTION

Low back pain (LBP) is the leading cause of disability worldwide,¹ and its prevalence is rising.² Nearly one-third of all older adults are estimated to have LBP,³ and the related societal costs are substantial. Within a 10-year period, there has been a nearly 400% increase in LBP-related Medicare charges associated with LBP management⁴; downstream costs related to progressive disability must also be considered. Several large, longitudinal studies have demonstrated that geriatric LBP is independently associated with a steeper rate of decline in performance-based mobility function (ie, gait speed, chair rise performance, balance).^{5,6} Further, poor mobility function among older adults is predictive of increased risk for mortality, institutionalization, and disability within 3 years.^{7,8} To mitigate the societal impact of LBP among older adults, effective interventions that address both pain and functional decline must be developed; however, existing barriers must be overcome to make progress on this front.

The National Institute on Aging has been keenly aware of the significant knowledge gaps preventing forward progress in this field of study and made a specific call for the “development and evaluation of therapeutic algorithms for comprehensive management of chronic pain specifically targeting the needs and challenges of older adults (eg, physical therapy).”⁹ Although there is long-standing expert consensus that patients with LBP should be classified into clinically relevant subgroups prior to treatment,^{10,11} no empirically derived treatment-based classification systems exist for older adults with LBP. There is strong evidence that classification approaches, where treatments are matched to the clinical impairments of a specific subgroup, can improve clinical outcomes for younger patients undergoing physical therapy treatment for LBP.¹²⁻¹⁴ To develop effective interventions, appropriate classification systems that account for the unique features of older adults with chronic LBP (ie,

degenerative joint changes, sarcopenia,⁵ maladaptive pain processing,¹⁵ functional limitations⁶) are needed.

One potential clinically relevant subgroup that has been discussed in the literature is individuals with chronic LBP and coexisting hip impairments. Offierski and MacNab initially defined hip-spine syndrome as the manifestation of low back symptoms because of an abnormal biomechanical relationship between the hip joint and the lumbar spine, likely related to degenerative pathology in 1 or both joints.¹⁶ Since then, several studies have identified relationships between hip impairments and LBP,¹⁷⁻²² but the hip-spine phenomenon has never been studied systematically or adequately addressed from a rehabilitative intervention standpoint. Thus, the 2 main study objectives were to investigate whether clinically relevant subgroups of older adults with chronic LBP could be identified on the basis of the presence of potentially modifiable hip impairments, using latent variable mixture modeling (LVMM) and to examine the relationship between these subgroups and key outcomes, including pain and performance-based mobility function, over a period of 12 months. Identification of homogeneous, clinically relevant subgroups based on modifiable hip impairments may inform targeted pain management approaches.

[H1]Methods

[H2]Study Overview

Individuals who were 60 to 85 years old were recruited from the greater Delaware area from March 2013 to December 2016 for this prospective cohort study. Individuals were included if they reported “moderate” LBP intensity ($\geq 4/10$) that occurred on most days per week (≥ 4 days) for ≥ 3 months. Individuals were excluded if they had nonmechanical LBP symptoms (eg,

unrelenting night pain, pain-related fever), severely limited mobility (eg, required more than a cane for household ambulation), a progressive neurological disorder (eg, Parkinson disease, multiple sclerosis), or a terminal illness. Additionally, individuals were excluded if they scored <18/30 points on the Mini-Mental State Examination²³; although a score of >24 is a generally accepted cutoff for normal cognition,²⁴ prior work has shown that individuals with lower education levels or from a racial minority group can score lower on this test without having cognitive impairment.²³ Participants were recruited from the community via advertisements placed in local newspapers, mailers, retirement communities, senior centers, and health fairs. Of the 432 individuals screened, 250 participants were included (Fig. 1).

Standardized clinical examinations (ie, comprehensive mobility function, low back, and hip assessments) occurred at baseline, 3 months, 6 months, and 12 months in a clinical research laboratory. Two licensed physical therapists, masked with regard to each other's assessments, performed each evaluation: 1 physical therapist completed the mobility function and low back assessments, whereas the other completed the hip assessment. This study was approved by the University of Delaware Institutional Review Board, and the protocol aligned with principles outlined in the Helsinki Declaration. Each participant provided written informed consent prior to enrollment.

[H2]Sample Characteristics

Participants reported their age, sex, race, employment status, education level, the duration of their LBP symptoms, and medication-use across various pharmaceutical classes (ie, opioids,

antidepressants, anxiolytics, and anticonvulsants). Height and weight were measured using a digital scale, and body mass index (BMI) was calculated. The 15-item Geriatric Depression Scale (GDS-15) was used to measure depressive symptoms; higher scores indicate more symptoms,²⁵ and a cutoff value of ≥ 5 is used to detect clinical depression.²⁶ The Cumulative Illness Rating Scale (CIRS) was used to quantify comorbidity burden.²⁷ The CIRS is a 13-item scale that assesses the presence and severity of impairments in different bodily regions (eg, cardiac, respiratory, gastrointestinal, musculoskeletal); 0 = no impairment and 4 = extremely severe impairment, and higher overall scores indicate a greater comorbid disease burden. To complete the CIRS, the examiner performed a comprehensive medical history interview with the participant and asked additional questions for clarification as needed.²⁷ Participants also completed the Fear Avoidance Beliefs Questionnaire for physical activity, a measure of pain-related fear of movement; lower scores indicate less fear.²⁸

[H2] Hip Impairment Domain Measures

[H3] Hip symptoms. Specific items from the Hip Dysfunction and Osteoarthritis Outcome Score (HOOS),²⁹ an extension of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC),³⁰ were used to assess hip symptoms. We used HOOS items P4 to P8 to assess pain intensity, as they also constitute the WOMAC pain subscale; participants were asked about the amount of pain incurred during a variety of daily activities (eg, climbing stairs, walking, lying down). Similarly, we used HOOS items S4 and S5 to assess hip joint stiffness, as they also constitute the WOMAC stiffness subscale; participants were asked about the amount of joint stiffness experienced in the morning and later in the day. Scores for all items range from 0 to 4 (0

= none; 4 = extreme); higher scores indicate greater pain and/or stiffness. The psychometric properties of both the HOOS²⁹ and the WOMAC³⁰ have been established.

[H3]Hip strength. Hip strength was measured using a portable handheld dynamometer (Lafayette Instrument Co., Lafayette, IN, USA), which was calibrated annually. All strength tests were isometric, and standardized procedures were followed (Suppl. File 1). Three trials of hip flexion, abduction, adduction, internal rotation, external rotation, and extension strength measures were performed, and the highest value was recorded. Values were normalized to BMI; the weaker of the 2 sides (ie, right versus left lower extremity) was used to arrive at a single measure for these analyses. The reliability of these methods has been established and is reported in Supplemental File 1.³¹

[H3]Hip range of motion (ROM). Passive ROM was measured according to standardized procedures adapted from Reese and Bandy³² (Suppl. File 1). A goniometer was used to measure hip flexion, abduction, adduction, and extension ROM, and an inclinometer was used to measure hip internal and external rotation. The more restricted of the 2 sides was used for these analyses. The reliability of these methods has been established and is reported in Supplemental File 1.³³

[H3]Hip flexibility. The Modified Thomas Test, straight leg raise, and Ober Test were used to quantify hip flexor, hip extensor, and iliotibial band flexibility, respectively (Suppl. File 1). All measures were taken in degrees, and the more restricted of the 2 sides was used for these analyses. The reliability of these tests has been established and is reported in Supplemental File 1.^{32,34,35}

[H2]LBP Intensity and Disability

The pain thermometer, a reliable and valid measure among older adults,³⁶ was used to assess “best” and “worst” pain intensity in the last 24 hours as well as “current” pain intensity. Scores ranged from 0 to 10 (0 = no pain; 10 = worst possible pain); these ratings were averaged to form a composite LBP intensity rating, which is more consistent with recommendations made by the NIH Task Force on Research Standards for Chronic Low Back Pain.³⁷

The Quebec LBP Disability Questionnaire was used to measure LBP-related disability. Participants were asked how difficult it was to perform certain activities of daily living (eg, standing, walking) because of their LBP. For each of the 20 items, scores range from 0 to 5 (0 = not difficult at all to do; 5 = unable to do). Total scores range from 0 to 100, with higher scores suggesting greater LBP-related disability. The reliability and validity of the Quebec LBP Disability Questionnaire have been established among older adults.³⁸

[H2]Performance-Based Mobility Function

Self-selected gait speed, a strong predictor of poor health outcomes among older adults,³⁹ was measured with the GaitMat II System (EQ Inc, Chalfont, PA, USA), a 4-m-long electronic walkway that measures a variety of spatiotemporal gait characteristics. Participants were instructed to walk at their usual, comfortable speed; after 3 practice trials, participants performed 3 recorded trials. The average gait speed of the 3 recorded trials was recorded. The GaitMat II System has been found to have excellent reliability.⁴⁰

Six-minute walk test distance (6-MWD), a reliable and valid measure of functional endurance in older adults, was measured over a circular track as described previously.^{41,42} Participants were instructed to cover as much ground as possible in 6 minutes and were provided with standard verbal encouragement every 30 seconds.

[H2]Data Analyses

Mixture model. Mplus⁴³ version 7.31 was used to conduct the LVMM, allowing for the identification of potentially clinically distinct subgroups on the basis of baseline data from the hip impairment domains (ie, symptoms, strength, ROM, and flexibility). Models with 2 through 4 classes were examined, with the final number being determined by the Akaike information criterion, the Bayesian information criterion (BIC), the sample size–adjusted BIC, and entropy statistics. Lower scores for the Akaike information criterion, the BIC, and the sample size–adjusted BIC are desirable, whereas higher entropy values indicate better model fit. In addition, Vuong-Lo-Mendell-Rubin (VLMR) and adjusted VLMR likelihood ratio tests were used to determine whether a model with k classes was a better fit than a model with $k - 1$ classes. Significant values (ie, $P \leq .050$) indicated that a model with k class fit better than a model with 1 class less (ie, $k - 1$).

[H3]Subgroup characteristics. The resulting classes, or subgroups, were compared on demographic/health characteristics (eg, age, sex, BMI) and hip impairment domain measures. χ^2 tests were used for dichotomous measures, and analysis of variance or Kruskal-Wallis tests were used for between-group comparisons of continuous measures. Findings for continuous measures

were identical regardless of which test (ie, parametric or nonparametric) was used; results presented are from the analysis of variance models.

[H3]Longitudinal comparisons. Mixed models were used to compare the subgroups longitudinally over 4 time points: baseline, 3 months, 6 months, and 12 months. Primary outcomes for the mixed models were composite LBP intensity, Quebec LBP Disability Questionnaire scores, self-selected gait speed, and 6-MWD. Fixed effects included subgroup membership, time, and the subgroup \times time interaction as well as the covariates of age, sex, education level, employment status, BMI, GDS-15 scores, and CIRS scores. Because the time measurement occasions were unequally spaced (ie, baseline, 3 months, 6 months, 12 months), analyses were performed by modeling time as both unequally spaced and equidistant. Findings were consistent between both approaches; models with equally spaced time are presented for ease of interpretation. A compound symmetric heterogeneous covariance error was chosen by examining optimal model fit using the Akaike information criterion and the BIC.

[H2]Role of the Funding Source

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under award number R01AG0412202. The funder played no role in the design, conduct, or reporting of this study.

[H1]Results

[H2]Mixture Model Fit

Various numbers of classes were tested in LVMM (Tab. 1). A model with 3 classes exhibited the lowest BIC value; VLMR and adjusted VLMR *P* values indicated that a 3-class model was a significant improvement (*P* = .01) from a 2-class model and that a 4-class model did not significantly improve model fit. Therefore, for parsimony, a 3-class model was chosen. Average class probability for individuals in the most likely class was high: .983, .964, and .967 for classes 1, 2, and 3 (henceforth referred to as “subgroups”), respectively.

[H2]Subgroup Characteristics

Table 2 illustrates the differences between the 3 subgroups on hip impairment domain measures included in the LVMM (ie, hip symptoms, strength, ROM, and flexibility). Subgroups largely differed on the presence of hip symptoms and strength. Compared to subgroup 3, subgroups 1 and 2 had significantly lower symptom scores (ie, HOOS P4–P8 and HOOS S4 and S5) but did not differ from each other. Subgroup 1 was stronger across all hip strength measurements than subgroups 2 and 3; although subgroup 2 was stronger than subgroup 3 on several strength measurements, the differences were less pronounced. Therefore, we named the 3 clinically distinct subgroups as follows: subgroup 1 = strong and nonsymptomatic (SNS), subgroup 2 = weak and nonsymptomatic (WNS), and subgroup 3 = weak and symptomatic (WS).

Table 3 illustrates the differences in baseline demographic and health characteristics among subgroups: subgroups differed on several characteristics, with some exceptions (ie, Mini-Mental State Examination, presence of select health conditions, use of select medications, LBP duration, and Fear Avoidance Beliefs Questionnaire for physical activity scores). Compared with the WNS and WS subgroups, the SNS subgroup was younger, predominantly White, well-educated, and

not retired and had lower BMIs, fewer depressive symptoms, and fewer comorbidities.

Additionally, the SNS subgroup had a lower proportion of individuals with scores of ≥ 5 on the GDS-15 (SNS = 7.9%, WNS = 7.8%, and WS = 29.6%; $P \leq .001$), indicating a lower rate of clinical depression.

[H2]Longitudinal Comparisons

The covariates of age, sex, BMI, GDS-15 (ie, depressive symptoms), and CIRS (ie, comorbidity burden) had various associations with the 4 dependent variables (Supplemental File 2).

Generally, older age, female sex, higher BMIs, more depressive symptoms, and more comorbidities were predictive of higher composite LBP intensity, higher Quebec LBP Disability Questionnaire scores, and worse mobility performance.

The top panel of Figure 2 displays the adjusted class means for composite LBP intensity and Quebec LBP Disability Questionnaire outcomes over time. Models for both LBP outcomes followed similar patterns; main effects for subgroup ($P \leq .001$) and time ($P \leq .001$) were significant for both outcomes. SNS and WNS had lower pain intensity and disability scores than WS ($P \leq .001$ for all comparisons). Regardless of subgroup, composite LBP intensity and Quebec LBP Disability Questionnaire scores improved from baseline to 12 months ($P \leq .001$).

The bottom panel of Figure 2 displays the adjusted means for 2 performance-based mobility function outcomes over all time points. For self-selected walking speed, main effects for subgroup ($P = .01$) and time ($P = .01$) were significant. The SNS subgroup walked faster than the WNS ($P \leq .001$) and WS ($P = .01$) subgroups; there was no difference between the WNS

subgroup and the WS subgroup ($P = .17$). All subgroups improved from baseline to 12 months ($P \leq .001$). For 6-MWD, the subgroup \times time interaction effect was significant ($P = .02$); the SNS subgroup walked longer distances than both WNS and WS subgroups at baseline (both P s $\leq .001$) and significantly improved from baseline to 12 months ($P = .01$). The WS subgroup's 6-MWD did not change over time ($P = .29$), and was similar to the WNS subgroup's 6-MWD at baseline ($P = .99$). Although the WNS subgroup's 6-MWD initially increased, it regressed such that it was no different between baseline and 12 months ($P = .39$), nor was it different from the WNS subgroup's 6-MWD at 12 months ($P = .89$).

[H1]Discussion

Using an unsupervised statistical approach with minimal bias, we identified 3 clinically distinct subgroups, thus highlighting the unique role of hip impairments in characterizing older adults for whom the primary problem is chronic LBP. Our subgroup classification, which was established using LVMM, was based on the presence or absence of specific hip impairments, particularly strength and symptoms (ie, pain and stiffness). Identification of these subgroups allowed us to explore the longitudinal consequences of class membership; as expected, the subgroup with prevalent hip weakness and symptoms had the worst prognosis over 12 months. These findings suggest hip impairments may play an influential role in the clinical course of geriatric LBP; further subgroup exploration may allow for development and testing of tailored treatments for this population.

The 3 subgroups are primarily defined by hip symptoms and hip strength. Of note, hip ROM and flexibility measures did not play a substantive role in distinguishing subgroups. However, hip

symptoms were an important classification factor, which is unsurprising, given our previous findings: coexisting hip symptoms were more prevalent in older adults with chronic LBP compared to pain-free peers and were associated with worse function and health-related quality of life.²² The presence of movement-evoked hip pain, positional hip pain, and/or hip stiffness, which are all measured by the HOOS, may be indicative of hip joint pathology.⁴⁴ As for the role of hip strength in distinguishing LBP subgroups, existing literature supports the premise that altered hip muscle function is linked to LBP. In a recent systematic review, Sadler et al⁴⁵ concluded that LBP was associated with reduced gluteus medius muscle strength, particularly on the predominant side of LBP. Laboratory studies have also demonstrated poor endurance and coactivation of gluteus medius is predictive of LBP onset during prolonged standing among previously asymptomatic participants.^{46,47}

Given the strong role of hip symptoms and strength in distinguishing these subgroups, we named the subgroups based on the dominant impairment patterns seen in each group: SNS, WNS and WS. The SNS subgroup had hip strength values that were nearly twice that of the 2 “weak” subgroups, across all measures of hip strength. We acknowledge the lack of a known cut-point to definitively classify someone as strong versus weak; rather, we used the term *weak* in this context to denote there were 2 subgroups with a relatively lower strength level. Similarly, the WS subgroup had symptom scores approximately twice those of the nonsymptomatic subgroups. The consistent pattern in these 3 subgroups across hip impairment measures help to establish the meaningfulness of the distinct subgroups identified through VLMM.

Longitudinally, consistent patterns exist for each outcome across subgroups. From baseline, the WS

subgroup had worse pain outcomes with significantly higher pain intensity and Quebec LBP Disability Questionnaire scores (Fig. 2); although all 3 subgroups improved on pain outcomes over time, the between-group difference in pain remained stable. Of note, the difference in Quebec LBP Disability Questionnaire scores between the WS subgroup and the 2 nonsymptomatic subgroups (Quebec LBP Disability Questionnaire scores: WS = 38.8, WNS = 26.0, SNS = 21.6) exceeded the minimum detectable change value of 10.66 for the Quebec LBP Disability Questionnaire,³⁸ suggesting that the WS subgroup truly had greater LBP-related disability. From baseline, the WS subgroup also had worse mobility function, with significantly slower gait speed and shorter 6-MWD. A gait speed of <1.0 m/s indicates an increased risk for poor outcomes among older adults, including hospitalization and death⁴⁸; whereas, a gait speed of ≥ 1.2 m/s indicates exceptional life expectancy.⁴⁹ The gait speed of the 2 weak subgroups was ≤ 1.0 m/s at baseline and never reached the ideal gait speed of 1.2 m/s during follow-up; however, the SNS subgroup was not considered at risk at any point and reached the ideal gait speed by 6 months. For 6-MWD, all 3 subgroups had values in the normal range, but the SNS subgroup clearly outperformed the other groups. Further, only the SNS subgroup improved during follow-up. Notably, across all time points, the difference in 6-MWD between the SNS and the 2 weak subgroups exceeded the minimum detectable change of 58 m and the minimal clinically important difference of 50 m established for older adults who are healthy.⁵⁰ Collectively, these performance data strongly suggest meaningful between-group differences.

When considering the differences in patterns of pain versus functional mobility outcomes (Fig. 2), hip symptoms appear to be linked with worse longitudinal pain outcomes, whereas hip weakness is

more tied to worse performance-based function. For instance, the 2 nonsymptomatic subgroups cluster together in their pain-related longitudinal behavior and have better outcomes relative to the WS subgroup; but the 2 weak subgroups cluster together relative to function with worse outcomes over time. Given these patterns, it is logical that the LBP subgroup with both hip symptoms and weakness (ie, WS) would consistently have the worst outcomes. The consistency of longitudinal findings across multiple outcomes highlights the unique nature of each of these 3 subgroups.

Over the past decade, there has been a shift away from the consideration of physical impairments towards psychological factors in LBP management; this shift has been appropriate in younger populations where psychological factors have proven to be a predominant driver of chronic pain. However, older adults are vulnerable to age-related changes in all body systems, particularly the musculoskeletal system, and these senescent changes should be considered when developing age-appropriate treatment strategies.⁵¹ Our exploration of musculoskeletal hip impairments resulted in identifying 3 clinically relevant subgroups with different risk profiles; each subgroup likely requires a different focus to their LBP treatment approach.

A major advantage of our study design was the inclusion of potentially modifiable hip impairments needed to identify and name these subgroups. We postulate the subgroups identified herein could guide whether hip-focused interventions should augment traditional lumbar-focused approaches for older adults with chronic LBP. For instance, the SNS subgroup would not likely require any hip intervention, whereas the WNS subgroup might require hip strengthening. Finally, the WS subgroup would likely require an approach that addresses hip pain, stiffness, and weakness, in addition to targeted functional training. Given the substantial consequences (ie, falls, institutionalization,

mortality) associated with the combination of poor performance-based mobility function and multisite pain in older adults,^{7,8,52} this WS subgroup should be considered an at-risk group in need of special attention from the clinical community.

Although certain clinical characteristics were not relevant in distinguishing subgroups (eg, range of motion) or were not included in the model (eg, psychological factors), it does not necessarily mean such factors are unimportant relative to clinical outcomes. Impairments in these areas may contribute to worse physical function and greater disability, regardless of subgroup membership. Therefore, individual patients require comprehensive evaluations including, but also beyond, subgroup classification measures to develop effective treatment plans.

This study has several distinct strengths: the evaluation of community-dwelling older adults with LBP in the context of a prospective cohort design, with 12 months of follow-up data and a minimal attrition rate of 9%; *a priori* selection of reliable and valid hip impairment and performance-based outcome measures; masking of outcomes assessors to the results of the clinical hip assessment; and, the use of latent class analysis to identify distinct subgroups based on the presentation of hip impairments. Additionally, key demographic characteristics (eg, sex, BMI) are similar between our sample and other geriatric LBP study samples.^{53,54}

[H2]Study Limitations

There are also several limitations to consider. First, in the context of this study, the results have not been replicated in a new set of participants to validate the findings; this limitation is partially mitigated by the robust and consistent longitudinal findings across multiple outcomes, highlighting

the unique nature of each distinct hip-spine subgroup. Another factor which limits the generalizability of our findings, is the lack of full racial/ethnic representation for our sample; additionally, our sample appears to be more highly educated than a nationally-representative of older adults with back pain.⁵⁵ These results should be replicated in a population-based sample with adequate representation. We were also unable to establish an inception cohort in this population because of the challenge of identifying older adults in the primary care setting with a truly new onset of LBP. Other important pain factors, such as neuropathy, were not considered in our study. Finally, previous injuries and structural changes to the hip may impact movement patterns and, thus, pain and physical function in different ways; we did not measure factors such as these in our study.

[H1]Conclusion

Globally, LBP is a principal factor leading to disability among older adults,⁵⁰ and effective treatments that address pain and functional mobility consequences are limited. This study has taken the first step in developing a comprehensive therapeutic algorithm by identifying clinically relevant subgroups within the geriatric LBP population and delineating the longitudinal pain and functional profiles for these subgroups. Since these classification subgroups are based upon hip impairments that are modifiable through physical rehabilitation, we are well-positioned to develop interventions based on the defining features of each subgroup. The next steps involve the following: replicating these findings in larger population-based samples; examining other clinical characteristics (eg, psychological factors) that may be associated with subgroup membership; exploring whether other clinical characteristics are stronger predictors of clinical outcomes, compared to the presence of hip symptoms or muscle weakness themselves; developing and testing tailored treatment approaches for each subgroup.

Author Contributions

Concept/idea/research design: G.E. Hicks, R.T. Pohlig, P.C. Coyle, J.M. Sions, D.K. Weiner, J.M. Pugliese

Writing: G.E. Hicks, R.T. Pohlig, P.C. Coyle, J.M. Sions, J.M. Pugliese

Data collection: G.E. Hicks, P.C. Coyle, J.M. Sions, J.M. Pugliese, T.O. Velasco, V.A. O'Brien

Data analysis: G.E. Hicks, R.T. Pohlig, P.C. Coyle, V.A. O'Brien

Project management: G.E. Hicks, J.M. Sions, J.M. Pugliese, T.O. Velasco, V.A. O'Brien

Fund procurement: G.E. Hicks

Providing participants: G.E. Hicks, J.M. Sions

Providing facilities/equipment: G.E. Hicks

Providing institutional liaisons: G.E. Hicks

Clerical/secretarial support: P.C. Coyle, J.M. Sions, T.O. Velasco, V.A. O'Brien

Consultation (including review of manuscript before submitting): G.E. Hicks, P.C. Coyle, J.M. Sions, D.K. Weiner, J.M. Pugliese

Ethics Approval: This study was approved by the University of Delaware Institutional Review Board.

Funding: Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health (award no. R01AG0412202).

Disclosure

The authors completed the ICMJE Form for Disclosure of Potential Conflicts of Interest and reported no conflicts of interest.

[H1]REFERENCES

1. Vos T, Flaxman AD, Naghavi M et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(9859):2163-2196.
2. Hoy D, March L, Brooks P et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis*. 2014;73(6):968-974.
3. Docking RE, Fleming J, Brayne C et al. Epidemiology of back pain in older adults: prevalence and risk factors for back pain onset. *Rheumatology (Oxford)*. 2011;50(9):1645-1653.
4. Weiner DK, Kim, Y, Bonino, P, Wang, T. Low back pain in older adults: are we utilizing healthcare resources wisely? *Pain Medicine*. 2006;7(2):143.
5. Hicks GE, Simonsick EM, Harris TB et al. 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):1420-1424.

6. 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):793-797.
7. Guralnik JM, Simonsick EM, Ferrucci L et al. 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.
8. 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):556-561.
9. PA-09-193: Mechanisms, measurement, and management of pain in aging: from molecular to clinical (R01). Published 2009. Updated 2012. Accessed May 18, 2020. <https://grants.nih.gov/grants/guide/pa-files/pa-09-193.html>
10. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. *Phys Ther*. 1995;75(6):470-485; discussion 485-479.
11. Borkan JM, Koes B, Reis S, Cherkin DC. A report from the Second International Forum for Primary Care Research on low back pain. *Spine*. 1998;23(18):1992-1996.
12. Brennan GP, Fritz JM, Hunter SJ, Thackeray A, Delitto A, Erhard RE. Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: results of a randomized clinical trial. *Spine*. 2006;31(6):623-631.
13. Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a

- randomized clinical trial. *Spine (Phila Pa 1976)*. 2003;28(13):1363-1371; discussion 1372.
14. Long A, Donelson R, Fung T. Does it matter which exercise? A randomized control trial of exercise for low back pain. *Spine (Phila Pa 1976)*. 2004;29(23):2593-2602.
 15. Lautenbacher S. Experimental approaches in the study of pain in the elderly. *Pain Med*. 2012;13 Suppl 2:S44-50.
 16. Offierski CM, MacNab I. Hip-spine syndrome. *Spine (Phila Pa 1976)*. 1983;8(3):316-321.
 17. Clohisy JC, Knaus ER, Hunt DM, Leshner JM, Harris-Hayes M, Prather H. Clinical presentation of patients with symptomatic anterior hip impingement. *Clin Orthop Relat Res*. 2009;467(3):638-644.
 18. Sembrano JN, Polly DW, Jr. How often is low back pain not coming from the back? *Spine (Phila Pa 1976)*. 2009;34(1):E27-32.
 19. Ben-Galim P, Ben-Galim T, Rand N et al. Hip-spine syndrome: the effect of total hip replacement surgery on low back pain in severe osteoarthritis of the hip. *Spine (Phila Pa 1976)*. 2007;32(19):2099-2102.
 20. Hsieh PH, Chang Y, Chen DW, Lee MS, Shih HN, Ueng SW. Pain distribution and response to total hip arthroplasty: a prospective observational study in 113 patients with end-stage hip disease. *J Orthop Sci*. 2012;17(3):213-218.
 21. Chimenti PC, Drinkwater CJ, Li W, Lemay CA, Franklin PD, O'Keefe RJ. Factors associated with early improvement in low back pain after total hip arthroplasty: a multi-center prospective cohort analyses. *J Arthroplasty*. 2016;31(1):176-179.

22. Hicks GE, Sions JM, Velasco TO. Hip symptoms, physical performance, and health status in older adults with chronic low back pain: a preliminary investigation. *Arch Phys Med Rehabil.* 2018;99(7):1273-1278.
23. 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.
24. Lopez MN, Charter RA, Mostafavi B, Nibut LP, Smith WE. Psychometric properties of the Folstein Mini-Mental State Examination. *Assessment.* 2005;12(2):137-144.
25. Friedman B, Heisel MJ, Delavan RL. Psychometric properties of the 15-item Geriatric Depression Scale in functionally impaired, cognitively intact, community-dwelling elderly primary care patients. *J Am Geriatr Soc.* 2005;53(9):1570-1576.
26. Dias F, Teixeira AL, Guimarães HC et al. Accuracy of the 15-item Geriatric Depression Scale (GDS-15) in a community-dwelling oldest-old sample: the Pietà Study. *Trends Psychiatry Psychother.* 2017;39(4):276-279.
27. Hudon C, Fortin M, Vanasse A. Cumulative Illness Rating Scale was a reliable and valid index in a family practice context. *J Clin Epidemiol.* 2005;58(6):603-608.
28. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain.* 1993;52(2):157-168.
29. Nilsson A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American

- Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. *Arthritis Care Res (Hoboken)*. 2011;63(Suppl 11):S200-S207.
30. Sun Y, Sturmer T, Gunther KP, Brenner H. Reliability and validity of clinical outcome measurements of osteoarthritis of the hip and knee--a review of the literature. *Clin Rheumatol*. 1997;16(2):185-198.
 31. Thorborg K, Petersen J, Magnusson SP, Hölmich P. Clinical assessment of hip strength using a hand-held dynamometer is reliable. *Scand J Med Sci Sports*. 2010;20(3):493-501.
 32. Reese NB, Bandy WD, Yates C. *Joint range of motion and muscle length testing*. 2nd ed. St. Louis, Mo.: Saunders/Elsevier; 2010.
 33. Prather H, Harris-Hayes M, Hunt DM, Steger-May K, Mathew V, Clohisy JC. Reliability and agreement of hip range of motion and provocative physical examination tests in asymptomatic volunteers. *PM R*. 2010;2(10):888-895.
 34. Clapis PA, Davis SM, Davis RO. Reliability of inclinometer and goniometric measurements of hip extension flexibility using the modified Thomas test. *Physiother Theory Pract*. 2008;24(2):135-141.
 35. Rose MJ. The statistical analysis of the intra-observer repeatability of four clinical measurement techniques. *Physiotherapy*. 1991;77(2):89-91.
 36. Ware LJ, Herr KA, Booker SS et al. 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):475-482.
 37. Deyo RA, Dworkin SF, Amtmann D et al. Report of the NIH Task Force on Research Standards for Chronic Low Back Pain. *Phys Ther*. 2015;95(2):e1-e18.

38. 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.
39. Middleton A, Fritz SL, Lusardi M. Walking speed: the functional vital sign. *J Aging Phys Act.* 2015;23(2):314-322.
40. Brach JS, Perera S, Studenski S, Newman AB. The reliability and validity of measures of gait variability in community-dwelling older adults. *Arch Phys Med Rehabil.* 2008;89(12):2293-2296.
41. 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):837-841.
42. Rikli RE, Jones CJ. The reliability and validity of a 6-minute walk test as a measure of physical endurance in older adults. *J Aging Phys Act.* 1998;6:363-375.
43. Muthén LK, Muthén BO. *Mplus User's Guide*. Eighth ed. Los Angeles, CA: Muthén & Muthén; 1998-2017.
44. Altman R, Alarcón G, Appelrouth D et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum.* 1991;34(5):505-514.
45. Sadler S, Cassidy S, Peterson B, Spink M, Chuter V. Gluteus medius muscle function in people with and without low back pain: a systematic review. *BMC Musculoskelet Disord.* 2019;20(1):463.
46. Marshall PW, Patel H, Callaghan JP. Gluteus medius strength, endurance, and co-activation in the development of low back pain during prolonged standing. *Hum Mov Sci.* 2011;30(1):63-73.

47. Nelson-Wong E, Gregory DE, Winter DA, Callaghan JP. Gluteus medius muscle activation patterns as a predictor of low back pain during standing. *Clin Biomech (Bristol, Avon)*. 2008;23(5):545-553.
48. Cesari M, Kritchevsky SB, Penninx BW et al. Prognostic value of usual gait speed in well-functioning older people--results from the Health, Aging and Body Composition Study. *J Am Geriatr Soc*. 2005;53(10):1675-1680.
49. Studenski S, Perera S, Patel K et al. Gait speed and survival in older adults. *JAMA*. 2011;305(1):50-58.
50. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. *J Am Geriatr Soc*. 2006;54(5):743-749.
51. Simon CB, Hicks GE. Paradigm shift in geriatric low back pain management: integrating influences, experiences, and consequences. *Phys Ther*. 2018;98(5):434-446.
52. Butera KA, Roff SR, Buford TW, Cruz-Almeida Y. The impact of multisite pain on functional outcomes in older adults: biopsychosocial considerations. *J Pain Res*. 2019;12:1115-1125.
53. 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.
54. Makris UE, Paul TM, Holt NE et al. The relationship among neuromuscular impairments, chronic back pain, and mobility in older adults. *PM R*. 2016;8(8):738-747.

55. Rundell SD, Karmarkar A, Nash M, Patel KV. Associations of multiple chronic conditions with physical performance and falls among older adults with back pain: a longitudinal, population-based study. *Arch Phys Med Rehabil.* 2021.

TABLES:

Table 1.

Mixture Model Fit for Different Class Structures^a

No. of Classes	Model Fit Criteria				Model Fit Improvement Tests	
	AIC	BIC	ssBIC	Entropy	VLMR <i>P</i>	aVLMR <i>P</i>
2	22,477.93	22,780.77	22,508.14	0.942	.01	.02
3	22,011.17	22,433.75	22,053.34	0.930	.01	.01
4	12,962.78	22,505.09	22,016.90	0.912	.76	.76

^aAIC = Akaike information criterion; aVLMR = adjusted Vuong-Lo-Mendell-Rubin (VLMR) likelihood ratio test; BIC = Bayesian information criterion; ssBIC = sample size-adjusted BIC.

Table 2.

Between-Class Differences in Hip Impairment Measures Using Univariate ANOVA and Follow-Up Post Hoc Pairwise Comparisons^a

Measure	Mean (SD) for:			Model <i>P</i>	Post Hoc Pairwise Comparison <i>P</i>		
	Class 1 (n = 76)	Class 2 (n = 103)	Class 3 (n = 71)		Class 1 vs Class 2	Class 1 vs Class 3	Class 2 vs Class 3
Hip symptoms							
HOOS pain items (0–20)	3.2 (3.0)	2.9 (3.0)	6.9 (3.7)	<.001	.53	<.001	<.001
HOOS stiffness items (0–8)	1.7 (1.6)	1.4 (1.5)	3.3 (1.7)	<.001	.24	<.001	<.001
Hip strength (kg normalized to BMI)							
Flexion	0.65 (0.23)	0.38 (0.15)	0.33 (0.12)	<.001	<.001	<.001	.06
Abduction	0.37	0.21 (0.06)	0.18 (0.06)	<.001	<.001	<.001	.01

	(0.08)						
Adduction	0.41 (0.09)	0.22 (0.06)	0.18 (0.07)	<.001	<.001	<.001	.01
Internal rotation	0.34 (0.08)	0.17 (0.05)	0.14 (0.05)	<.001	<.001	<.001	.01
External rotation	0.42 (0.11)	0.22 (0.07)	0.18 (0.07)	<.001	<.001	<.001	.01
Extension	0.35 (0.13)	0.18 (0.08)	0.14 (0.07)	<.001	<.001	<.001	.01
Hip ROM (°)							
Flexion	93.4 (7.3)	93.2 (2.9)	90.2 (8.2)	.02	.96	.02	.01
Abduction	23.5 (5.5)	23.9 (6.7)	22.6 (7.7)	.43			
Adduction	13.4 (6.0)	12.5 (7.1)	12.2 (6.31)	.47			
Internal rotation	21.0 (9.8)	24.8 (10.8)	24.1 (10.6)	.05	.02	.08	.65
External rotation	31.8 (9.9)	30.0 (10.3)	27.7 (9.6)	.05	.26	.02	.13
Extension	4.4 (5.3)	4.3 (5.9)	4.1 (5.9)	.95			
Hip flexibility (°)							
Modified Thomas Test	0.3 (7.9)	0.9 (6.3)	-0.5 (7.1)	.45			
Straight leg raise test	54.2 (12.1)	55.7 (14.0)	53.9 (13.7)	.64			

Ober Test	4.0 (7.1)	4.0 (8.0)	4.6 (7.2)	.83			
-----------	-----------	-----------	-----------	-----	--	--	--

^aANOVA = analysis of variance; BMI = body mass index; HOOS = Hip Dysfunction and Osteoarthritis Outcome Score; ROM = range of motion.

Table 3.

Between-Class Differences in Baseline Demographic and Health Characteristics Using

Univariate ANOVA or χ^2 Tests^a

Characteristic	Full Sample (n = 250)	Class 1, SNS (n = 76)	Class 2, WNS (n = 103)	Class 3, WS (n = 71)	P
Age, y, mean (SD)	69.7 (6.8)	67.4 (6.1)	71.4 (6.8)	69.7 (7.0)	<.001
Sex, women	128 (51.2)	14 (18.4)	62 (60.2)	52 (73.2)	<.001
Race, White	213 (85.2)	70 (92.1)	91 (88.3)	52 (73.2)	.01
Employment status, retired	182 (72.8)	47 (61.8)	81 (78.6)	54 (76.1)	.03
Education, college graduate	145 (58.0)	51 (67.1)	64 (62.1)	30 (42.3)	.01
MMSE score of ≥ 24	247 (98.8)	76 (100)	101 (98.1)	70 (98.6)	.49
BMI, kg/m ² , mean (SD)	29.4 (5.7)	27.2 (4.5)	29.9 (5.4)	31.3 (6.5)	<.001
GDS-15 score, 0–15, mean (SD)	2.2 (2.7)	1.7 (2.4)	1.7 (2.2)	3.4 (3.4)	<.001
Heart disease	31 (12.4)	10 (13.2)	9 (8.7)	12 (16.9)	.25
Stroke	10 (4.0)	1 (1.3)	7 (6.8)	2 (2.8)	.15
Diabetes	41 (16.4)	11 (14.5)	14 (13.6)	16 (22.5)	.25
Osteoarthritis	131 (52.4)	35 (46.1)	53 (41.4)	43 (60.6)	.21
Visual impairment	95 (38.0)	30 (39.5)	36 (35.0)	29 (40.8)	.70
Hearing impairment	79 (31.6)	27 (35.5)	31 (30.1)	21 (29.6)	.68
CIRS score, 0–52, mean (SD)	9.4 (3.8)	8.1 (3.4)	9.5 (3.5)	10.7 (4.1)	<.001

Opioid use	18 (7.2)	3 (3.9)	7 (6.8)	8 (11.3)	.21
Antidepressant use	50 (20.0)	11 (14.4)	22 (21.4)	17 (23.9)	.31
Anxiolytic use	20 (8.0)	3 (3.9)	10 (9.7)	7 (9.9)	.29
Anticonvulsant use	25 (10.0)	4 (5.3)	12 (11.7)	9 (12.7)	.24
LBP duration, y, mean (SD)	5.7 (7.7)	5.8 (7.6)	5.3 (7.2)	6.1 (8.7)	.80
FABQ-PA score, 0–24, mean (SD)	13.5 (6.2)	13.0 (6.0)	13.0 (6.2)	14.9 (6.2)	.11
Composite LBP level, 0–10, mean (SD)	3.1 (1.5)	2.9 (1.4)	2.7 (1.4)	3.8 (1.5)	<.001
Quebec LBP Disability Questionnaire score, 0–100, mean (SD)	28.3 (16.6)	21.6 (12.6)	26.0 (16.4)	38.8 (15.8)	<.001
Self-selected gait speed, m/s, mean (SD)	1.01 (0.23)	1.14 (0.19)	0.98 (0.22)	0.92 (0.23)	<.001
6-MWD, m, mean (SD)	488 (133)	589 (78)	453 (133)	429 (119)	<.001

^aData are reported as number (percentage) of participants unless otherwise indicated. ANOVA = analysis of variance; BMI = body mass index; CIRS = Cumulative Illness Rating Scale; FABQ-PA = Fear Avoidance Beliefs Questionnaire for physical activity; GDS-15 = 15-item Geriatric Depression Scale; LBP = low back pain; MMSE = Mini-Mental State Examination; 6-MWD = 6-min walk test distance; SNS = strong, no symptoms; WNS = weak, no symptoms; WS = weak, symptoms.

FIGURE LEGENDS

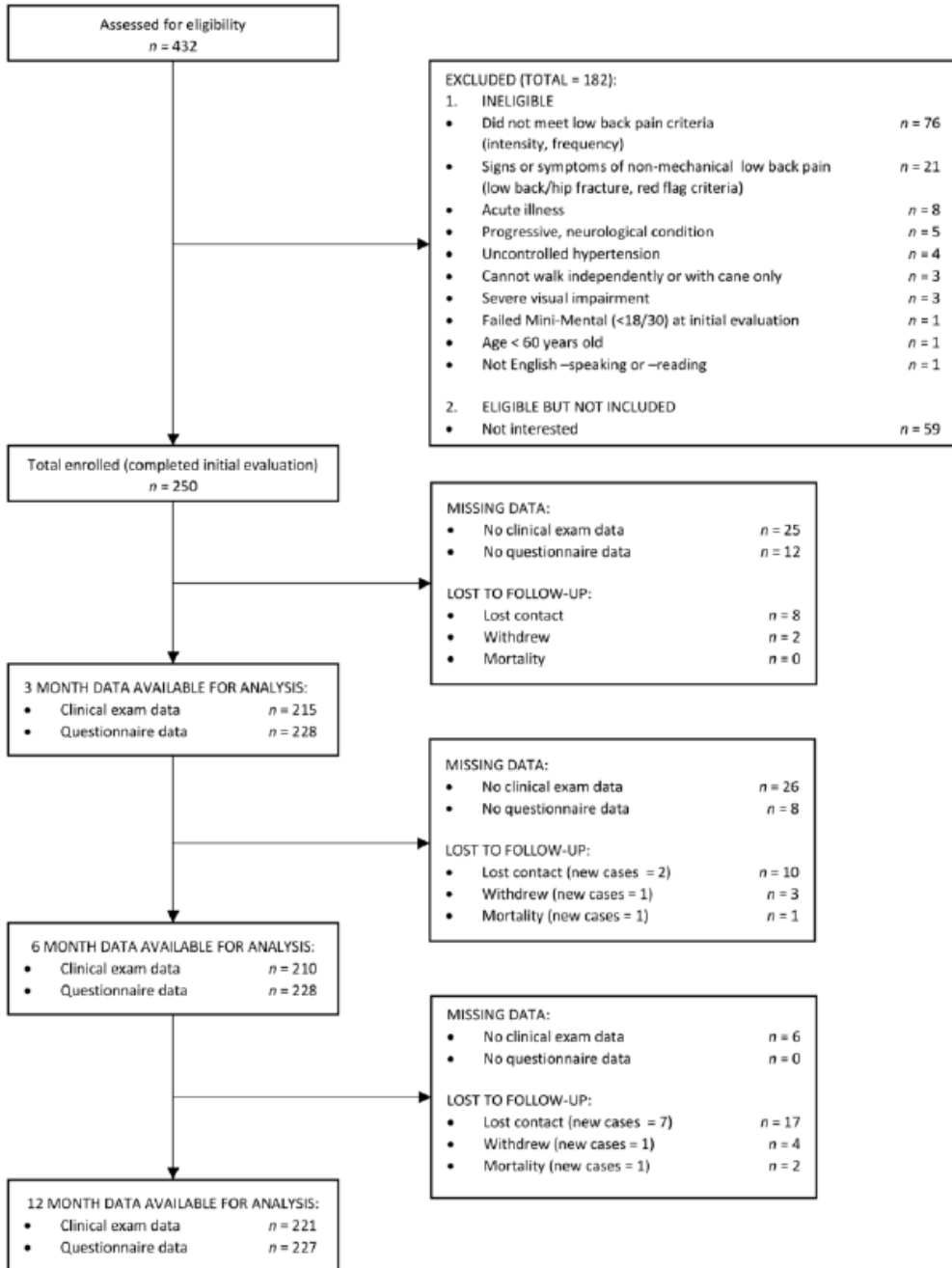


Figure 1. Study flow diagram.

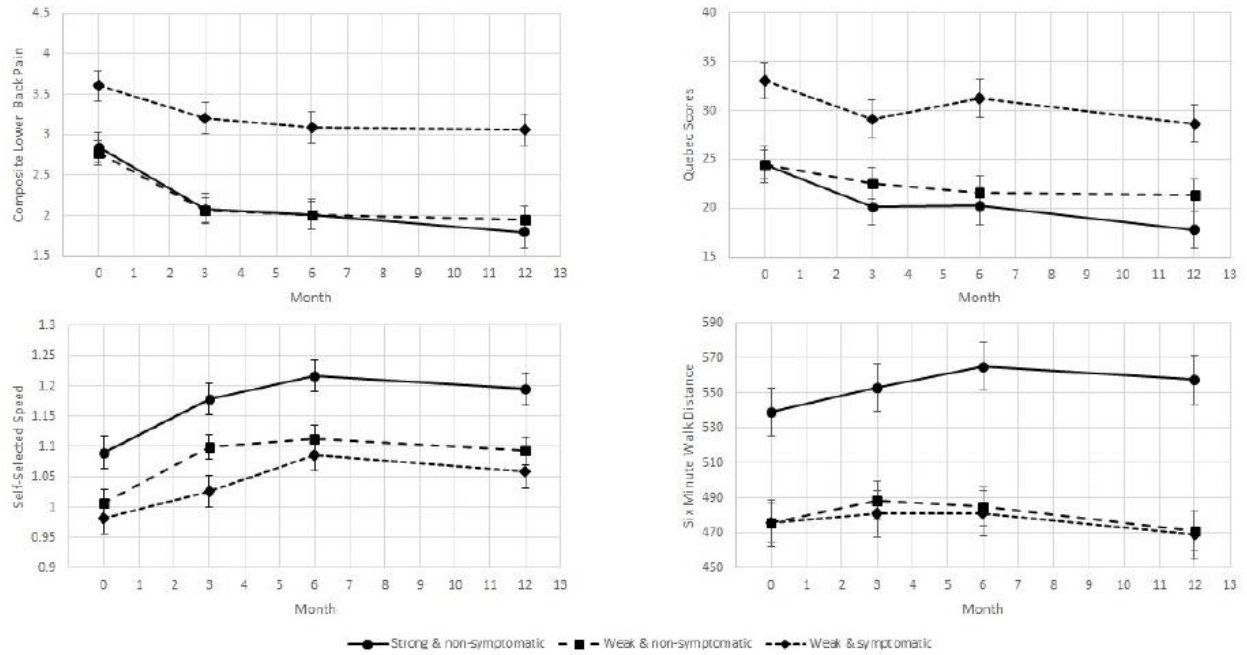


Figure 2. Class means for pain and mobility function outcomes across all time points, adjusted for age and body mass index.