Study Title: Determining the Impact of a Physiotherapist-Led Primary Care Model for Low Back Pain: Study Protocol and Analysis Plan for a Cluster Randomized Controlled Trial and Embedded Process Evaluation

Authors:
Jordan Miller[a,b], jordan.miller@queensu.ca
Catherine Donnelly[a,b], catherine.donnelly@queensu.ca
Chad McClintock[a,b], cwm2@queensu.ca
Kevin Varette[a,b], varettek@queensu.ca
Yeimi Camargo [a,b], yeimi.c@queensu.ca
Jacquelyn Marsh[c], jmarsh2@uwo.ca
Monica Taljaard[d,e], mtaljaard@ohri.ca
Mir Sanim Al Mamun[a], mirsanimal.mamun@queensu.ca
Geneviève Bacchus[a], 20gkb1@queensu.ca
David Barber[f], barberd@queensu.ca
Lynn Cooper[g], lkcooperbes@rogers.com
Simon French[a,h], simon.french@mq.edu.au
Michael Green[b,f,i], mgreen@nosm.ca
Jonathan Hill[j], j.hill@keele.ac.uk
Joy MacDermid[c], jmacderm@uwo.ca
Kathleen Norman[a], kn4@queensu.ca
Julie Richardson[k], jrichard@mcmaster.ca
Joan Tranmer[b,l], tranmerj@queensu.ca
Timothy Wideman[m], timothy.wideman@mcgill.ca

[a]School of Rehabilitation Therapy, Queen's University, Louise D Acton Building, 31 George Street, Kingston, Ontario, Canada, K7L 3N6
[b]Health Services and Policy Research Institute, Queen's University, Abramsky Hall, 21 Arch Street, K7L 3N6
[c]School of Physical Therapy, Western University, Elborn College, 1201 Western Rd, London, Ontario, Canada, N6G 1H1
[d]Ottawa Hospital Research Institute, Ottawa Hospital, Civic Campus, 1053 Carling Ave, Ottawa, Ontario, Canada, K1Y 4E9
[e]Epidemiology and Public Health, Univerisity of Ottawa, 75 Laurier Avenue East, Ottawa, Ontario, Canada, K1N 6N5
[f]Department of Family Medicine, Queen's University, 220 Bagot Street, Kingston, Ontario, Canada, K7L 3G2
[g]Canadian Injured Workers Alliance, 127 Ravenwood Avenue, Thunder Bay, Ontario, P7B 4H7
[h]Department of Chiropractic, Macquarie University, Level 2/75 Talavera Road, Sydney, New South Wales 2109, Australia
[i]Northern Ontario School of Medicine, 955 Oliver Road, Thunder Bay, Ontario P7B 5E1
[j]School of Allied Health Professionals, Keele University, MacKay Building, Newcastle-under-Lyme, United Kingdom, ST5 5BG
[k]School of Rehabilitation Science, McMaster University, Institute for Applied Health Sciences Building, 1400 Main Street West, Hamilton, Ontario, Canada, L8S 1C7
[l]School of Nursing, Queen's University, Cataraqui Building, 92 Barrie Street Kingston, Ontario K7L 3N6, Canada
[m]School of Physical and Occupational Therapy, McGill University, Hosmer House, 3630 prom Sir-William-Osler, Montreal, Quebec, Canada, H3G 1Y5

Corresponding author:
Jordan Miller, Associate Professor,
School of Rehabilitation Therapy, Health Services and Policy Research Institute, Queen's University,
Louise D Acton Building, 31 George Street, Kingston, Ontario, Canada, K7L 3N6
Jordan.miller@queensu.ca

Version Date: November 20, 2025

Trial Status

This trial was registered prospectively in January, 2020 at ClinicalTrials.gov (NCT04287413). Recruitment of primary care sites (clusters) was completed in June 2023 (after being delayed due to the COVID-19 Pandemic). Cluster randomization was performed in July 2023. Patient participant recruitment was initiated in October, 2023 and completed in November, 2024. Data collection is ongoing as of the submission of this protocol with data collection planned for completion by December 19, 2025. Data analysis will be carried out in accordance with this pre-specified analysis plan and will commence following the completion of data collection. No interim analyses have been conducted. Results are expected to be published in 2026.

Introduction

Low back pain (LBP) is the world's largest contributor to years lived with disability[1], costs the Canadian healthcare system between $6 and $12 billion annually[2], and is a leading contributor to lost work productivity[3,4]. The burden of LBP on the healthcare system is evidenced by frequent healthcare use, including unnecessary specialist consultations, diagnostic procedures[4,5], opioid prescriptions[6,7], and emergency department visits[8].

LBP is the fifth most common reason for physician visits[9-11] with primary care physicians being the most frequent first point of contact within the healthcare system for people with LBP[12]. A Lancet[13-16] series on LBP highlighted international expert consensus on the need to evaluate new primary care models that better support physicians, who receive limited training in the management of musculoskeletal (MSK) conditions[17,18] and report low confidence addressing LBP[19,20]. One of these models involves integrating physiotherapists (PTs) within primary care teams at the first point of contact, which has the potential to provide patients with a more focused LBP consultation, assist with evidence-based treatment delivery, reduce low-value care (such as inappropriate imaging, specialist physician referrals, and opioid prescriptions), and reduce primary care physician visits. If integrating PTs can reduce primary care physician visits, it could contribute to increasing the capacity of primary care teams to help address the growing challenge of providing universal access to primary care[21] in the context of more complex patient encounters[22,23], increasing prevalence of multiple chronic conditions[24], and an aging population[25].

Evidence suggests PTs can provide collaborative care[26-30] and implement recommendations from LBP primary care guidelines[31-34], including: screening for serious pathology, the need for diagnostics[35-37] and identifying risk factors for poor recovery[38-40]; providing reassurance, encouragement for early return to work, and exercise and physical activity recommendations[41]; and delivering targeted, psychologically informed interventions for those at risk of prolonged recovery[40]. Evidence from outside of Canada suggests that early guideline adherent PT care for LBP improves function and disability[42,43] while reducing diagnostic imaging, opioid prescriptions, and physician specialist referrals[43-45], and reduces per-person healthcare costs[46,47].

Evidence from observational research suggests that involving PTs in providing first-point-of-contact care (i.e., PT-led care) for those with work-related injuries in the US military resulted in workers being more satisfied with their care, receiving faster access to treatment, having fewer sickness absences, and using PTs and specialist physicians more appropriately[36,48-52]. Observational studies from the UK National Health Service on first-contact PT models of care indicate similar health outcomes, high levels of patient and physician satisfaction, shorter physician wait times, fewer work absences and diagnostic images, lower prescription medication use, and reduced costs[53-59].

The absence of high-quality randomized trials of PT-led primary care models leaves important unanswered questions about the process and impact of integrating PTs within primary care teams for people with LBP[60]. Specifically, there is a need for higher quality evidence on the impact of PT-led primary care on patient-oriented outcomes (e.g., function, pain, quality of life), health system outcomes (e.g., healthcare access, physician workload, emergency department visits, specialist physician referrals, medication use, diagnostic imaging), and societal outcomes (e.g., missed work, cost-effectiveness). Also, it is unclear how PTs will navigate primary care challenges, such as providing care for people presenting with multiple health concerns or addressing requests for medications, diagnostic imaging, or notes for work absences. This cluster randomized trial and embedded process evaluation will address these gaps by assessing the impact and implementation of a PT-led primary care model for LBP. The

results will inform primary care transformations across multiple health systems and potentially improve outcomes for patients with LBP.

The purpose of publishing this protocol and analysis plan is to transparently report our design and methods, and to transparently communicate our analytic plan in advance of carrying out our planned analysis to reduce the risk of analytic or reporting bias.

Research Questions:
1) Is a PT-led primary care model for LBP effective at improving disability (primary outcome), pain intensity, quality of life, global rating of change, patient satisfaction, and adverse events compared to usual physician-led primary care, when evaluated over a one-year period?
2) What is the impact of a PT-led primary care model for LBP on the health system and society (healthcare access, primary care physician workload, healthcare utilization, missed work, cost-effectiveness), evaluated over a one-year period?

Methods

Trial Design:

The trial is a parallel arm cluster randomized controlled trial at 20 primary care sites randomized 1:1 to a PT-led or usual physician-led primary care model for LBP. Randomization of practices, rather than patient participants, allows evaluation of a model where PTs are able to fully integrate within the primary care team and reduces potential contamination between study arms[61]. This protocol and analysis plan includes all of the items included in the Standard Protocol Items: Recommendations for Interventional Trial (SPIRIT) reporting guidelines[62,63]. See Table 1 for a summary of the schedule of enrolment, interventions and assessments.

Table 1. Schedule of enrolment, interventions, and assessments.

	STUDY PERIOD							
	Enrolment of clusters	**Allocation of clusters**	**Enrolment of patient participants**	**Post-allocation**				**Close-out**
TIMEPOINT	**Pre-intervention**	**Pre-intervention**	**0 (baseline)**	***6 wk***	***12 wk***	***6 mo.***	***9 mo.***	***12 mo.***
ENROLMENT OF PRIMARY CARE SITES:								
Eligibility screen	X							
Informed consent	X							
Allocation to study group		X						
ENROLMENT OF PATIENT PARTICIPANTS:								
Eligibility screen			X					
Informed consent			X					
INTERVENTIONS:								
PT-led primary care			◆—	—	—	—	—	—◆

Usual care			←					→
ASSESSMENTS:								
Participant characteristics and demographic information			X					
Participant health outcomes (RMDQ, pain intensity, EQ-5D-5L, PSEQ, PCS, TSK-11, PHQ-2)			X	X	X	X	X	X
Participant health and experience outcomes (Satisfaction, GROC, adverse events)				X	X	X	X	X
Health system and societal outcomes (provider encounters, EHR utilization, healthcare utilization, missed occupational activities, assistance required)				X	X	X	X	X
Implementation measures (timely access, PT as first point of contact, access to PT, LBP management)			X	X	X	X	X	X
Implementation measures (patient adherence)				X				
QUALITATIVE INVERVIEWS:								
Patient participants					X			
Primary care providers and leads, medical office assistants*							X	X

* These interviews started 9 months after study launch.

The trial incorporates a multi-methods process evaluation informed by United Kingdom Medical Research Council (MRC) guidance for developing and evaluating complex interventions[64,65], the Consolidated Framework for Implementation Research (CFIR)[66-69], and findings from the process evaluation from our pilot trial[70]. Process evaluations assess how interventions were implemented and under what conditions. They are important to help interpret trial results (e.g., explaining why an intervention fails or has unexpected consequences, or why it works and how it can be refined). Process evaluations can be particularly valuable for informing future implementation and sustainability of complex healthcare interventions[65,71]. Our process evaluation is informed by and intended to refine our program theory[64,72] for the PT-led primary care model program for LBP (**Figure 1**). The program theory

describes how the PT-led primary care model for low back pain is intended to lead to improvements in trial outcomes.

Allocation of Participating Primary Care Teams to Trial Groups:
Restricted randomization is recommended for cluster randomized trials to ensure the arms are balanced at baseline. We used covariate-constrained randomization[73] with 1:1 ratio to the intervention and comparison arms to retain the merits of random allocation while ensuring baseline balance across the arms. We stratified by location (Southeastern Ontario or Interior British Columbia) and included the following covariates: number of active patients and rural vs. urban setting. Our maximum tolerable difference in the rurality indicator was 1 and the strata balancing criteria for number of active patients was set to 10%. An independent statistician implemented the procedure using a SAS macro[74]. Concealment was maintained by ensuring each practice had an anonymized code and performing randomization after all sites were recruited.

Methods for Protecting Against Sources of Bias:
Due to the trial design and interventions being compared, the PTs, patient participants, other members of the primary care team, and research assistants have not been blinded. While limitations due to the inability to blind are unavoidable, we have taken additional steps to minimize the risk of bias suggested for cluster trials[75]. A common challenge in cluster trials is identification and recruitment bias when patient recruitment must take place after cluster allocation[76]. To minimize the risks of these biases, we invited consecutive patients to participate, provided the same information about the trial to both groups prior to participants consenting (not revealing the cluster allocation until after consent and baseline data collection), employed research assistants unfamiliar with the patients to recruit and consent participants, and rigorously applied inclusion/exclusion criteria at all sites[77].

Inclusion/Exclusion Criteria:
For primary care sites, inclusion criteria were: ≥2 family physicians and/or nurse practitioners and ≥ 1,500 active patients. Exclusion criteria were: already having a PT on the team or no space available for a PT to practice. Patient participants were screened prior to consent by a trained research assistant at both intervention and comparison sites. Inclusion criteria for patients participants were: adults (≥19 years of age) with LBP of any duration, who were able to read, write and speak English. Exclusion criteria were: known cancer that could possibly contribute to LBP or inability to complete the scheduled follow-ups over one-year.

Recruitment:
We recruited 20 sites (14 in Ontario and 6 in British Columbia, Canada). Site recruitment focused primarily on the engagement of contacted sites to reduce the risk of site withdrawl[78]. For patient participant recruitment, staff at each primary care site screened patients for their willingness to participate when they called to book an appointment. If patients booked online, their reason for visit was screened by the clinic staff and staff reached out to potential participants for their willingness to be contacted by a research assistant. Trained research assistants screened potential participants using inclusion/exclusion criteria and enrolled consenting patients. Patients who raised LBP with any provider during a clinical encounter were also identified as potential participants. In these cases, they were directed to the administrative team, who then connected them with a research assistant for screening and enrollment. Balance in recruitment across arms and across clusters was monitored throughout the study and strategies were implemented to maintain consistent implementation of recruitment processes across sites.

Sample Size and Power Calculations:

We used the methodology of Teerenstra et al.[79] to calculate the required number of clusters based on an ANCOVA analysis for the primary outcome (Roland-Morris Disability Questionnaire [RMDQ]) at 12 months adjusting for baseline. Our target sample was 20 clusters (10 per arm) allowing for 1560 patient participants (78 per practice before attrition). Conservatively accounting for attrition of two clusters (resulting in 18 clusters) and 20% of patients (resulting in 63 patients/cluster), this target would achieve 90% power to detect a minimally important mean difference of 2.5 points[80] (Cohen's d=0.4) using a two-sided α=0.05 and assuming a standard deviation (SD) of 5.7 points based on pilot study data, a conservative intracluster correlation coefficient of 0.1[81], a cluster autocorrelation coefficient (correlation between cluster means at baseline and follow-up) of 0.5, and an individual autocorrelation coefficient (correlation between participant scores at baseline and follow-up) of 0.6 informed by our pilot study[82]. An average rate of 1.5 patient participants per site each week over one year would achieve the recruitment target of 1560 participants, which we anticipated based on our pilot study where we recruited and retained four sites and achieved a recruitment rate of >1.7 participants/site/week. Recruitment of 400 patient participants across 18 clusters (22 participants/cluster) would achieve 80% power.

Clinical Partner Sites:
This study is being coordinated from a primary Research Coordinating Centre at Queen's University with 20 participating primary care sites (clusters) between Southeastern Ontario (14) and Interior British Columbia (6), Canada. Sites include representation from urban and rural settings in both regions to facilitate generalizability.

Patient and Public Involvement:
A senior advisor for the Ontario Ministry of Health contributed to the plan for this study. An individual with lived experience is a co-author on this protocol (LC) who has been engaged in the design of this study, including the identification and decision on the trial outcome measures that are important to patients. Their involvement in carrying out the study includes pre-testing of data collection methods, interpretation of process evaluation results, and participation in the development of knowledge mobilization products (i.e., designing summaries appropriate for patient organizations).

Trial Interventions:
1. Index Intervention: PT-Led Primary Care Model for LBP
The index intervention involves integrating a PT within the primary care team and making them available at the first point of contact for people with LBP. Patients with LBP are encouraged to book with the PT except when the primary reason for their visit is medication renewal or when additional health concerns require another provider's attention. The intervention has four components: 1) assessment and screening; 2) brief individualized intervention; 3) supporting patients to access appropriate health services based on assessment findings; and 4) providing additional PT care to people with an unmet need. Ten registered PTs, who have completed two days of training on this care model, are participating in delivery of the model across the 10 practices randomized to the PT-led primary care arm.

Assessment and Screening: The assessment and screening includes: taking a history; screening for pathology (e.g., cauda equina syndrome, traumatic fracture, cancer); physical and neurological examination; application of evidence-based tools to identify comorbid health conditions (e.g., 2-item Patient Health Questionnaire[83] for depression) requiring additional care; and using a validated tool (Keele STarT Back Tool[39,40]) to identify risk factors associated with persistent LBP and disability.

Brief Individualized Intervention: The PTs provide a brief individualized intervention at the initial visit. This intervention is intended to be based on primary care guidelines for LBP [34] and consists of effective communication to validate the patient's experiences[84] and allow the patient to disclose the impact of their LBP on their lives[85,86], cognitive reassurance[87], individually tailored exercises[88,89], and advice/strategies to stay active[90].

Supporting Patients to Access Appropriate Health Services Based on Assessment Findings: It is intended that the PT collaborates with the patient to identify appropriate health services based on assessment findings, collaborates with relevant primary care team members to provide the needed care, and integrates health service providers from outside of the team considering the patient's needs and access to health services. First, the PT identifies potential pathology needing urgent referral (e.g., cauda equina, fracture, infection). Next, they identify comorbid conditions that require collaboration with other primary care team members (e.g., physicians, nurse practitioners, nurses, social workers, occupational therapists, pharmacists, and dieticians). For example, people who screen positive for depression are referred to their physician or nurse practitioner or a member of the mental health team. Finally, primary care PTs refer to community PT as informed by the patient's score on the Keele STarT Back tool[39,40] (if appropriate). The STarT Back tool categorizes patients with LBP into low, medium, or high risk of persistent pain and disability based on physical and psychosocial risk factors[39]. The recommended matched treatment for low-risk patients is to provide self-management advice and to avoid referral and investigations where possible. This intervention is brief and it is expected that it can take place at the first visit. The recommended matched treatment for medium-risk patients is referral for standard community PT, and the matched treatment for high-risk patients is PT care from the primary care PT who received specific training aimed at reducing physical and psychosocial risk factors for chronic pain and disability as part of the two-day training[91,92]. This stratified approach to care has demonstrated improved function, quality of life, and cost-effectiveness in comparison to usual care in the UK[40]. When a need for PT care is identified, the primary care PT helps these patients navigate the available PT resources (i.e., private PT clinics when the patient has private health insurance; government funded PT for those who meet the criteria; or PT in primary care for those without access to a PT elsewhere).

Providing Additional PT Care to People with an Unmet Need: The PT provides additional care at the primary care site for patients who are appropriate for PT (based on STarT Back score) and who have barriers to accessing PT care in their community. For example, people who may benefit from physiotherapy but do not have access to services—due to a lack of private or government insurance coverage or because of geographic or transportation barriers—are offered additional physiotherapy care. Care includes evidence-based, guideline consistent management, such as individualized education[93] exercise[88,89] behavioural approaches[91] To avoid duplication of available PT services, patient participants with private or government funded health insurance for PT are referred to external PT services.

2. Comparison: Usual Physician-Led Care

The usual physician- or nurse practitioner-led primary care intervention has been unstandardized to reflect usual primary care clinical practice in Canada. Our pilot study suggested the most common management approaches in the usual care model included: diagnostic imaging (12% of patients), medication renewal (21%), new medication prescription (14%), notes to employers (12%), and referral (PT 23%, chronic pain clinic 5%, massage therapy 5%, physician specialists 5%, and dietician 2%). Interventions provided or recommended by the physician have been and will continue to be recorded and monitored throughout the full trial.

There are no required or prohibited concomitant treatment options for either treatment arm. Should a participant choose to explore other or additional treatment options within their primary care team or from other health service providers, they are free to do so. This was made clear in the letter of information and consent form provided to the participants upon entry to the study.

Duration of Treatment Period:
The PT-led primary care model is being implemented over a one-year period from the time of consent. All participants were offered an initial assessment with the PT and have access to the PT for follow-up needs for a one-year period after their initial assessment. The majority of participants classified as low-risk using the STarT Back tool and those with private or government-funded health insurance for PT are intended to only see the PT in primary care at their initial visit, but will have access to the PT as a member of their primary care team throughout the one-year follow-up period if they need or want a follow-up visit. Participants identified as medium or high risk without access to PT elsewhere are offered additional PT care from the PT in primary care. The frequency and duration of the treatment plan is determined by the PT and patient participant. In our pilot study, participants who took part in ongoing care received an average of four visits over eight weeks.

Intervention Modifications:
We do not anticipate encountering a situation that would require the withdrawal of a participant for safety concerns related to the PT-led primary care or usual physician-led primary care. As per normal primary care and PT practice for LBP, the intervention will be modified by the primary care team to maintain participant safety as part of the model of care (e.g., modification of exercise in response to increases in pain that may be experienced with an exercise intervention, discontinuation of medications if adverse effects are experienced). Due to the low-risk nature of the study and the fact that patient participants maintained access to their usual primary care providers, a Data Monitoring Committee was not established. Ongoing oversight and patient participant safety monitoring was provided by the study investigators, including monitoring all adverse events reported at all time points.

Data Collection and Management:
All patient participant-reported outcome measures are being collected from participants using REDCap (Research Electronic Data Capture), a secure online survey and data capture tool hosted at Queen's University[94,95]. A distinct link for the surveys is being sent to each participant at each assessment time-point. We are providing the option of completing the questionnaires in-person or by phone if participants have barriers to completing them online. Data are also being extracted from participants' electronic health records (EHR) related to care provided for LBP. Data are being extracted using pre-piloted EHR extraction forms.

At study completion, responses to the surveys will be exported directly from REDCap to encrypted and password protected datasets and stored securely in Microsoft OneDrive at Queen's University. All data collected from the EHR, along with a master linking log that links study identification numbers with participants, will be stored in password-protected and encrypted files in OneDrive. The linking log will be permanently destroyed at the end of the data analysis period. Qualitative interview recordings will be transcribed, deidentified, and stored securely in OneDrive.

We implemented multiple strategies to promote participant retention across all time points. Research assistants sent reminders every 2-3 days through personalized emails, phone calls, and text messages to encourage timely survey completion. When requested, surveys were completed in-person or by phone to strengthen engagement and minimize loss to follow-up.

Frequency and Duration of Follow-Up:
Patient characteristics and demographic information were collected at baseline. All patient-reported outcome measures are being collected at baseline, 6 weeks, and 3, 6, 9, and 12 months from the initial visit, with the primary comparison at 12 months. Patient satisfaction, global rating of change, adverse effects, and healthcare utilization are being collected at follow-up time points only. Patient adherence to PT recommendations is being collected at the 6-week follow-up.

Participant Characteristics:
The following characteristics and demographic information were collected from patients at baseline: age, first three digits of their postal code (to determine rural/urban status), biological sex, gender, identification as a member of a racialized group, duration of the current episode of LBP, previous history of LBP, number of other pain locations, highest level of education achieved, household income, and work status. In addition, the following questionnaires were administered at baseline as potential covariates and to inform subgroup analyses as part of our process evaluation.

Functional Comorbidity Index: an 18-item list of comorbidities that are associated with physical functioning. Each comorbidity is assigned a score of 1 and the total score is the sum of the comorbidity elements[96-98].

Keele STarT Back Tool: categorizes patients with LBP into low, medium, or high risk of persistent pain and disability based on physical and psychosocial risk factors[39,40,91].

Patient Health and Experience Outcomes:
Self-Reported Disability (Primary Outcome): using the RMDQ, which demonstrates reliability, validity, and responsiveness in people with acute and chronic LBP[99,100].

Pain Intensity: using a Numeric Pain Rating Scale (0-10)[101] for pain at rest, pain when walking, and pain when lifting a bag of groceries from the floor. Each will be reported on a scale of 0 (no pain at all) to 10 (worst imaginable pain).

Health-Related Quality of Life: using the EuroQOL-5D (EQ-5D-5L)[102], which demonstrates good reliability and validity, and is suitable for economic evaluation in LBP[103]. The EQ-5D-5L VAS score (0-100) will be reported as a patient health outcome. The EQ5D responses will be converted to an EQ5D index value using the value set calculated for the Canadian context[104]. The index value will be reported as a patient-level health outcome and will also be used to calculate quality-adjusted life years (QALY) for our outcome in the economic evaluation.

Global Rating of Change: using an 11-point global rating of change scale, from a great deal better (+5) to a great deal worse (-5), as has been recommended for self-reported rating of change[105,106].
Participant satisfaction: using an 11-point scale for satisfaction with care, extremely dissatisfied (-5) to extremely satisfied (+5).

Adverse events: using an adverse events questionnaire consistent with reporting guidelines[107,108] that asks: 1) if the patient participant has experienced any adverse event(s) as a result of any of the treatments received; 2) what adverse event(s) were experienced; 3) how long the event(s) lasted; and 4) how severe each adverse event was. For analysis, adverse events will be identified as serious or non-serious. An adverse event will be identified as serious if any of the following criteria are met: the participant requires in-patient hospitalization or an emergency department visit due to the adverse event,

the adverse event results in significant and persistent disability (beyond 72 hours), or the adverse event is life-threatening or results in death. These responses are being monitored as they are completed in order to provide ongoing oversight of patient participant safety.

The following measures will be assessed as secondary outcomes, reported in the trial results, and included in the process evaluation as potential mechanisms through which the PT-led primary care model influences LBP-related disability.

Self-efficacy: confidence in abilities to participate in usual activities using the Pain Self-Efficacy Questionnaire (PSEQ)[109,110]

Psychosocial risk factors for persistent pain and disability: The Pain Catastrophizing Scale (PCS) [111,112], Tampa Scale of Kinesiophobia 11 (TSK-11) [113,114], and 2-item Patient Health Questionnaire (PHQ-2) [83,115] will measure psychosocial factors associated with pain-related disability.

Health System and Societal Outcomes:
Primary care physician or nurse practitioner encounters: the number of new and repeat primary care physician or nurse practitioner visits for LBP per patient. This measure will be considered a proxy for a potential increase in primary care team capacity achieved if the LBP-related workload of primary care physicians or nurse practitioners is reduced, thus increasing their availability to provide care to other patients.

Healthcare utilization within the primary care team: using data being collected from the EHR: consultations with other primary care team members (e.g., physicians, nurse practitioners, nurses, social workers, and occupational therapists) and group programming accessed within the primary care organization.

Healthcare utilization outside of the primary care team: using self-report data from follow-up surveys cross checked with reports the electronic health record when possible: medications used; walk-in clinic visits; ED visits; inpatient hospital stays; surgeries, injections, and other interventional procedures; visits to specialist physicians; diagnostic imaging; and visits to other health professionals outside the primary care team (e.g., chiropractors, massage therapists, occupational therapists, physiotherapists, chronic pain clinics).

Missed Occupational Activities: using self-report data from follow-up surveys: time (days) lost from paid employment, volunteer, homemaking, or educational activities related to LBP.
Assistance required: using self-report data from follow-up surveys: paid and unpaid assistance required. For example, self-care (e.g., taking medications, dressing/undressing, going to the bathroom, bathing/showering, grooming), shopping/groceries, meal preparation, housework, managing finances, or transportation (e.g., to a medical appointment).

Costs: Total costs per person will be calculated by summing direct healthcare costs and indirect costs using a human capital approach for missed occupational activities. Sources of direct healthcare cost data: Intervention costs will include the PT salary and training needed to carry out the intervention. Costs for publicly funded healthcare services will be obtained from the Ontario Ministry of Health Schedule of Benefits[116]. Medication costs will be obtained from the Ontario Drug Benefit formulary. Expenses related to health services funded by private insurance or out of pocket will be collected through self-report at all follow-up assessments. Other costs incurred by the participant related to their LBP are also being collected by self-report, including support or assistance for self-care, housework,

shopping, or transportation (e.g., to healthcare appointments). The total direct costs will be determined by multiplying the quantity of resource use by the corresponding unit cost, summing the total cost over each follow-up interval, and then calculating the mean cost at each follow-up time point, as well as an overall mean cost for the entire study period. *Indirect costs:* Non-healthcare costs will be limited to loss of productivity using a human capital approach[117]. The mean provincial wage reported by Statistics Canada will be used to assign a monetary value to time lost from paid employment. The minimum wage value in Ontario and BC will be used to place a value on time lost by those who were retired and time lost from volunteer, homemaking, caregiving, or educational activities.
Cost effectiveness: We will conduct a cost-utility analysis from both societal (primary) and health payer (secondary) perspectives to meet the needs of all knowledge users. For both societal and health payer perspectives, we will estimate the incremental cost-per-QALY gained[118-120].

Implementation Measures:
The following measures will be used to assess how the PT-led primary care model and usual physician- or nurse practitioner-led primary care model were implemented.
Timely access to LBP care: using the percentage of patient participants with LBP who are assessed within 48 hours of calling for an appointment. Participants who learned about the study and were invited to participate during an appointment for their LBP will not be included in this analysis.
PT as the first point of contact: using the percentage of patient participants with LBP in the PT-led primary care arm who visited a PT as their first point of contact for the current episode of LBP.
Access to PT services: using the percentage of patients who are classified as medium or high risk on the STarT Back screening tool who access PT (as endorsed by guidelines[121]).

LBP management: using the following data being collected from the EHR to describe the LBP management provided to patient participants in each arm: education; exercise; psychological approaches; referrals to other primary care team members (e.g., primary care physicians, nurse practitioners, nurses, social workers, occupational therapists, group programming); referrals made to health professionals outside of the primary care team; medications prescribed, deprescribed, and suggested; diagnostic imaging ordered; lab work ordered and received; notes to employers or insurers; interprofessional communications with the primary care team; and other interventions provided.
Patient adherence to recommendations: adherence to recommendations from the primary care PT is being collected at the 6-week follow-up. For patient participants identified as medium- or high-risk using the STarT Back classification, whether or not they accessed recommended PT (either through a referral to a community PT or through the PT in primary care) is being collected as part of our health utilization survey questions.

Qualitative Interviews:
Semi-structured qualitative interviews are being conducted with patients, PTs, other health professional primary care team members, medical office assistants, and primary care organizational leaders. Interviews with patient participants, PTs, and other health professionals have two goals: 1) to explore the experiences and perspectives with the PT-led primary care model for LBP; and 2) to understand how the model of care was implemented, how the intervention interacted with its context, and barriers/facilitators to implementation. Interview guides for interviews with patient participant s, PTs, and other health professionals start by exploring experiences with the PT-led primary care model for LBP using open-ended questions and probing. The interview guides then focus on asking participants about how the model of care was implemented, how the intervention interacted with its context, and barriers/facilitators to implementation using questions constructed to align with the CFIR domains[67,68,122]. The interview guides for medical office assistants and primary care organizational

leaders will focus on how the model was implemented and contextual factors influencing implementation using the CFIR. Interview guides were pre-piloted with persons with lived experience as patients and primary care team members prior to conducting the interviews.
We are using purposive sampling to recruit 8-12 patient participants with diversity in terms of age, gender, race, household income, LBP-related disability, LBP duration, STarT Back risk categories, and primary care site. We are inviting all PTs who are involved in implementing the PT-led primary care model for LBP to participate in an interview. We are purposively sampling 10-15 primary care health professionals who work with a PT in the PT-led primary care model for LBP, ensuring variation in terms of professional background, gender, and primary care site. We are using purposive sampling to identify four to eight medical office assistants and four to eight primary care organizational leaders who have experienced implementation of the PT-led primary care model. The concept of information power[123] related to our study objectives is being used to determine when to stop interviewing within each informant group based.

Patient participants were asked during their initial consent process for the main trial whether they were willing to be contacted for a follow-up qualitative interview about their experiences with the PT-led model of care. Willing patient participants were contacted within 12 weeks of enrollment. Based on our purposive sampling criteria, research assistants contacted potential participants who had agreed to be contacted to provide additional details about the interview purpose, review the consent process, and arrange a convenient time for the interview. Participants who agreed to take part were then sent a letter of information and consent form in advance. At the time of the interview, the research assistant confirms that the participant has reviewed the consent form, responds to any questions, and obtains verbal consent before proceeding. Recruitment of healthcare providers, medical office assistants, and organizational leaders is being carried out by a study coordinator and research assistant who are familiar with the participating clinical teams. Potential non-patient participants are being invited via email or in-person. Those who indicate interest are scheduled for an interview. As with patient participants, they will receive the letter of information and consent form ahead of the interview, and verbal consent is obtained at the start of the session after confirming their understanding and answering any questions.

Protocol and Analysis Plan Amendments:
Changes to the protocol and analysis plan will be communicated by amending the trial registry at ClinicalTrials.gov and reported in the full trial publication. Investigators and participants will be communicated with as appropriate based on the changes.

Analysis

Effectiveness analysis
All quantitative analyses will be conducted as per the Intention-To-Treat principle. Descriptive statistics for baseline characteristics and primary and secondary outcomes will be reported by arm using means (SD) or medians (interquartile range) for continuous variables and count (percent) for categorical variables. All analyses will be performed in SAS, version 9.4 (SAS Institute Inc; Cary, NC). Differences between arms will be compared, accounting for site clustering using linear mixed models and generalized estimating equations (GEE), and significance will be reported with p-values.

<u>Patient health and experience outcomes</u>
For our patient health outcomes, we will use linear mixed models to estimate individual patient participant outcomes adjusting for clustering by primary care centre. Our primary outcome (RMDQ) with repeated measures at baseline and the 6-week, 3-, 6-, 9-, and 12-month follow-up time points will be analyzed using linear mixed regression (using *PROC MIXED* in SAS), estimated using restricted

maximum likelihood (REML) estimation and a Kenward-Rogers degrees of freedom correction[124] to account for a small number of clusters. The model will include fixed effects for time, intervention group by time interaction, (omitting the group main effect to ensure baseline differences are constrained to zero[125]), factors used in the covariate-constrained allocation procedure[126] (rurality of the cluster, number of active patients) and other pre-specified covariates associated with LBP-related disability (patient participant age[127], sex[128], income[129], highest level of education achieved, duration of current episode of LBP[130], Functional Comorbidity Index score[127]). The correlation in repeated measures on the same participant will be modeled using a suitable covariance structure, identified using information criteria (AIC/BIC). To account for clustering within practices, site will be modeled as a random effect. The intervention effect will be obtained as the adjusted least square mean difference between arms at 12 months, with 95% confidence intervals. Secondary comparisons will be obtained using least square mean differences at intermediate time points.

The use of REML estimation under an assumption of Missing At Random (MAR) allows the use of all available data without the need for multiple imputation. To examine the risk of bias due to missing data, we will compare the characteristics of those remaining and those lost to follow-up to identify factors associated with attrition. We will perform a sensitivity analysis for a missing not at random (MNAR) departure from our MAR assumption using a delta-adjusted imputation pattern mixture model[131,132] approach to investigate the robustness of our trial outcomes with regard to the missing values of the RMDQ. Within this sensitivity analysis approach, we will start with the posterior distributions suggested by an imputation model using multiple imputation by chained equation. Our imputation model will incorporate all variables in our primary analysis model, along with the last-observed-before-time covariates, and additional covariates needed for the pattern mixture. The sensitivity parameter (delta) will be introduced to explore how a departure from MAR affects results by specifying a maximum delta for each pattern of missing. We will set the maximum delta to be twice the residual sample standard deviation from the observed data fit of the primary linear mixed model. Our sensitivity analysis will use multiple imputation 9 times to generate estimated treatment effects for a range of sensitivity[133].

Pain intensity, health-related quality of life (EQ-5D-5L), PSEQ, PCS, TSK-11, and PHQ-2 outcomes will be analyzed as described for the primary outcome, adjusted for the same covariates. Patient satisfaction and global rating of change outcomes have no baseline measures, and will be analyzed using ordinal logistic regression with clusters as random effects, adjusting for the same covariates as described above. When individual items are missing from within any of the questionnaires, we will use simple mean imputation as suggested by Chavance[133]. Serious adverse events will be presented descriptively by arm due to low expected counts. Any adverse events (yes/no) will be presented as incidence rates with confidence intervals and compared by calculating relative risks with confidence intervals from robust Poisson regression using GEE-type robust variance estimators (using *PROC GLIMMIX* with *EMPIRICAL* option in SAS) to account for clustering[134] and using an exchangeable working correlation matrix. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for small number of clusters in all models comparing incidence rates[135]. In the case of non-convergence or unstable estimates due to the small number of clusters, we would attempt to fit the model using an independent working correlation matrix. Other information related to non-serious adverse events (i.e., severity and duration) will be reported descriptively.

We have planned a secondary analysis, a responder analysis[136], to compare the proportion of participants who experience a meaningful improvement in disability (RMDQ) in the PT-led primary care model arm versus the usual care arm. We will define a meaningful improvement as an improvement of greater than or equal to 30% improvement on the RMDQ, corresponding to an established minimally important difference among people with LBP[137,138]. We will calculate the proportion of participants who

experience a meaningful improvement in each arm and compare between groups using robust Poisson regression, with GEE-type robust variance estimators to account for clustering[130]. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for small number of clusters[135].

Healthcare utilization and missed occupational activity outcomes

For healthcare utilization and missed occupational activities outcomes, we will estimate the average effect at the patient level across the population (marginal models), accounting for clustering using GEE-type robust variance estimators and robust Poisson regression with an exchangeable working correlation matrix. All of the models generated for our healthcare utilization and missed occupational activities outcomes will include the same covariates as our analysis for patient health outcomes: age, sex, highest level of education achieved, income, duration of current episode of LBP, and Functional Comorbidity Index score (individual level), and primary care site rurality and number of active patients (cluster level). In the case of non-convergence or model instability, a possibility for any binary outcomes with very low or high event rates given our small number of clusters[139], we will attempt to model outcomes using an independence working correlation matrix. If we continue to experience issues with non-convergence or instability with alternate covariance structures, we will remove covariates, starting with duration of pain and income based on theoretical grounds and existing evidence on the strength of relationships between our covariates and our outcomes.

Primary care physician or nurse practitioner visits will be presented as rates and compared using rate ratios with adjusted Poisson or negative binomial regression. Other healthcare utilization within the primary care team (whether there were any other consultations with interprofessional team members and whether or not there was group programming accessed), and healthcare utilization outside of the primary care team (medications, diagnostic imaging, walk-in clinic visits, ED visits, specialist physician visits, emergency department visits, hospital admissions, interventional procedures, surgeries, other health provider visits) will be presented as incidence rates and compared by calculating relative risks with confidence intervals using robust Poisson regression, accounting for clustering[134]. These models will use time as an offset to account for variable follow-up times and will incorporate empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for small number of clusters[135,140,141]. Time (days) lost from occupational activities (paid employment, volunteer, homemaking, or educational activities) and assistance required (hours of paid assistance, hours of unpaid assistance) due to LBP will be presented as rates and compared using rate ratios with negative binomial regression.

Economic evaluation

We will estimate the cost-effectiveness of the PT-led care model from both a societal (primary) and health system payer (secondary) perspective to meet the needs of all knowledge users. The total costs will be determined by multiplying the quantity of resource use (or lost days) by the corresponding unit cost (or hourly wage), summing the total cost over each follow-up interval, and then calculating the mean cost at each follow-up time point, as well as an overall mean cost for the entire one-year study period. Results will be presented as aggregated and disaggregated costs. Utility data will be generated using EQ-5D-5L index values (ie. utility scores) from all follow-up assessment time points. We will estimate QALYs for every participant, using the area under the curve approach, assuming linear interpolation between the measurements. To accommodate the hierarchical structure of the data, we will use bivariate multilevel modeling to estimate the incremental cost-effectiveness ratio using a calculation of cost-per-QALY gained for PT-led primary care versus usual care[118-120] and to estimate the incremental net benefit at various willingness to pay values. To account for clustering of study sites, as well as heterogeneity in costs and treatment effect across jurisdictions, we will model treatment group as

a fixed effect and the study site a random effect; adjusting for the same covariates as the primary analyses as fixed effects in our models. We will conduct a probabilistic sensitivity analysis using 10,000 Monte Carlo simulations to present the uncertainty in our cost-effectiveness estimates. Simulation results will be plotted on cost-effectiveness planes and we will generate cost-effectiveness acceptability curves to display the probability that the PT-led care model is cost-effective across a range of willingness to pay thresholds.

Subgroup analyses

In alignment with sex and gender equity in research (SAGER) guidelines[142], we plan to conduct exploratory analyses for each of our effectiveness outcomes for males and females to explore potential sex differences in each of these outcomes. We will include sex and its interaction with time and group by time in the models. We will report the interaction p-value along with forest plots to visualize the subgroup treatment effects, along with 95% confidence intervals.

Process evaluation analysis

The multi-methods process evaluation analysis will assess how the PT-led primary care model for LBP was implemented, the potential mechanisms of the model, the experiences and perspectives of patients and primary care team members toward the model, and how the context influenced implementation and outcomes. The analysis is guided by and intended to inform refinements to our program theory[64,72] for the PT-led primary care model program for LBP (**Figure 1**).

We will assess how the PT-led primary care model and usual care model were implemented and how the model influenced healthcare for people with LBP by:

1) Describing and comparing the proportion of patient participants who received timely access to LBP care between trial arms. The intended implementation outcome is that patient participants seeking primary care for LBP receive timely access (within 48 hours) to a LBP assessment from a PT in the PT-led primary care model for LBP. Access within 48 hours will be presented as incidence rates with confidence intervals and compared by calculating relative risks with confidence intervals using GEE-type variance estimators to account for clustering[134], using an exchangeable working correlation matrix. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for small number of clusters[131-133]. We will incorporate the same covariates as with our effectiveness analysis.

2) Describing the proportion of patient participants in the PT-led primary care arm who saw the PT as the first point of contact for the current episode of LBP.

3) Describing and comparing the proportion of patient participants who are categorized as medium or high risk on the STarT Back tool who access PT services between trial arms. The intended implementation outcome is that patient participants with LBP at medium or high risk of ongoing pain and disability access PT services, either through a referral (for participants who have access to PT services) or through additional PT care from the primary care PT (for participants who would not otherwise have access to PT services). We will report the proportion of participants at medium or high risk receiving PT care as incidence rates with confidence intervals and compared by calculating relative risks with confidence intervals using GEE-type variance estimators, accounting for clustering[134], using an exchangeable working correlation matrix. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for small number of clusters[131-133]. We will incorporate the same covariates as with our effectiveness analysis.

4) Describing and comparing the LBP management provided between trial arms and comparing the care provided to practice guidelines using recent World Health Organization (WHO) LBP guidelines[143] to assess the alignment of the care provided with practice guideline recommendations. The 10 recommended interventions in the WHO guidelines are: structured education, exercise, needling

therapies, spinal manipulation, massage, operant therapy, cognitive behavioural therapy, non-steroidal anti-inflammatory drugs, topical cayenne pepper, and multicomponent biopsychosocial care. Interventions that are recommended against include: traction, ultrasound, transcutaneous electrical nerve stimulation, and lumbar supports. Additionally, referrals for diagnostic imaging or physician specialist visits for spinal injections or surgery consultations are rarely needed for LBP and thus will be compared between arms for a potential reduction in low-value care[144,145]. The intended implementation outcome is that patient participants receive care recommended in practice guidelines and do not receive care recommended against in practice guidelines, and that a low proportion of people receive *referrals or prescriptions* for diagnostic imaging, physician specialist, or medications outside of NSAIDs. Receipt of each intervention recommended by the guidelines, recommended against by the guidelines, and referrals (for imaging, spinal injections, or surgical consults) will be presented as incidence rates with confidence intervals and compared by calculating relative risks with confidence intervals using GEE-type robust variance estimates, accounting for clustering[134], using an exchangeable working correlation matrix. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for small number of clusters[135]. We will incorporate the same covariates as with our effectiveness analysis.
5) Describing adherence to PT primary care recommendations. The intended implementation outcome is that patient participants report actioning the initial recommendations for physical activity and exercise and that those who were recommended to access PT services report doing so.

To assess potential mechanisms of the PT-led primary care model, if there is a significant treatment effect, we will conduct a series of mediation analyses[146] to assess whether changes in LBP-related disability is explained by changes in self-efficacy (PSEQ) or changes in psychosocial risk factors (PHQ-2, PCS, TSK-11) for persistent LBP-related disability. We will carry out this mediation analysis for each of the self-efficacy and psychosocial risk factor variables in our entire sample and in the subgroup of patient participants identified as high-risk using the STarT Back classification. We hypothesize that change in self-efficacy will mediate the relationship between the intervention arm and RMDQ score in the full patient participant sample and that change in psychosocial risk factors will mediate the relationship between treatment arm and RMDQ score in the high-risk subgroup. We will use a stepwise approach to exploring the temporal trends and dynamics of the treatment effect across repeated measures as proposed by Beril and colleagues[147], along with our theoretical insights regarding the mediation effect, when considering the appropriate mediation model. The indirect effect, or the intervention effect that can be explained by the mediator, will be determined as the difference between the total effect of the intervention and the direct effect of the intervention[148,149]. The significance of this effect can be used to statistically evaluate the possibility of mediation[150,151]. In this context, the total effect is the intervention effect obtained in the full model described in the trial. To be consistent with our primary outcome measure in the trial, we will be concerned with the time-specific mediation on the RMDQ outcome at the 12-month follow-up[152,153]. These effect measures will allow us to present the proportion of the total effect that is mediated through the respective measures of interest[154]. Development of the causal/associated conceptual model has allowed us to consider, and control for where needed, mediation analysis assumptions; that is, there is no intervention-outcome, mediator-outcome, or intervention-mediator confounding or mediator-outcome confounding that is influenced by the intervention itself[149,155,156]. Additionally, as part of exploring potential mechanisms, we will describe the proportion of any differences in total costs that are due to differences in healthcare utilization costs (and if so what healthcare services) and what proportion of any differences in total costs are due to differences in costs due to missed occupational activities.

We will further explore potential mechanisms for intervention effects by conducting a planned subgroup analysis for all process outcomes, RMDQ, and costs based on STarT Back risk classification (low,

medium, or high), recognizing that this analysis is exploratory and likely to be underpowered. We hypothesize that the participants in the PT-led group classified as low risk will: be less likely to receive requisitions for diagnostic imaging, referrals to specialist physicians, and prescriptions for medications; demonstrate reduced costs; and show similar disability outcomes in comparison to patient participants in the usual care group. We hypothesize that participants in the PT-led group classified as medium risk will: be more likely to access PT services; receive LBP management closer aligned with guidelines; receive fewer requisitions for diagnostic imaging, prescriptions for medications, and referrals to specialist physicians; demonstrate reduced costs; and have reduced LBP-related disability. Finally, we hypothesize that participants in the PT-led group classified as high risk will: be more likely to access PT services; receive LBP management closer aligned with guidelines and more targeted care for psychosocial risk factors of persistent LBP-related disability; show greater reductions in the psychosocial risk factors; demonstrate reduced costs; and have reduced LBP-related disability.

To explore the experiences and perspectives of patients, PTs, and other primary care health professionals who have participated in the PT-led primary care model for LBP, qualitative interviews with patient participants, the PTs, and other health professional primary care team members will be recorded, transcribed, and coded independently by two investigators. An inductive qualitative analysis will be completed in an interpretive description tradition[157,158]. Interpretive description was chosen because of its emphasis on adding an interpretive lens during the analysis process in order identify meaningful themes from the data that can be applied in practice. tice. This qualitative approach is therefore well aligned with our process evaluation goals to understand experiences and perspectives with the PT-led primary care model, potentially leading to refinements in the model of care, the program theory, or plans for scaling the intervention if effective. To promote rigour, we will use: two independent coders for the first two to three transcripts for each group (patient participants, PTs, other health professionals) and meet to reach agreement on the initial coding structure; reflexivity journaling and reflexive dialogue amongst team members throughout the analytic process; field notes and written memos; prolonged engagement within the data; and an audit trail of the research process and analytic decisions[159-163].

To explore how the context influenced the implementation of the PT-led primary care model for LBP, a concurrent mixed methods analysis will be conducted using the quantitative data collected to describe how the intervention is being implemented (analysis described above) and qualitative interview data from patient participants, PTs, other primary care health professionals, medical office assistants, and primary care organizational leaders. Qualitative analysis will begin by having multiple research team members immerse themselves in the data. Data will be coded by two independent research team members using an in-depth deductive (codes derived from CFIR constructs[164]) and inductive (codes derived from the data) coding process in alignment with the CFIR User Guide[164] and previous research[165-167]. Further, we will code relationships between constructs to capture how constructs interact, and how implementation determinants relate to how the model was implemented. Integration of qualitative and quantitative will be achieved through the design (convergent mixed methods), methods (merging), and interpretation (narrative, and joint displays)[168]. Merging of qualitative and quantitative data will be achieved by linking participant characteristics, implementation data, and qualitative interview data. By bringing together the quantitative and qualitative data for analysis and linking this data at the level of the patient participant and cluster, we will be able to carry out an in-depth analysis of how context influenced implementation. Integration at the interpretation phase will be achieved using a weaving approach to interpret and report qualitative and quantitative findings together[168]. Joint displays will be used to bring data together in a visual display if appropriate.

Ethics and Dissemination:

Ethics approval for this study has been obtained from the Queen's University Health Science and Affiliated Teaching Hospitals Research Ethics Board (HSREB #6027847). Written consent was obtained from all participants willing to participate.

We plan to mobilize the knowledge generated through this cluster randomized trial and process evaluation through a series of peer-reviewed manuscripts, with the following foci: i) effectiveness of the PT-led primary care model for people with LBP (including patient health, and health system and societal outcomes); ii) how the PT-led primary care model for LBP was implemented; iii) potential mechanisms through which the PT-led primary care model for LBP influences patient health and health system outcomes (if the intervention is effective); iv) experiences and perspectives of patients with LBP who participated in the PT-led primary care model; (v) experiences and perspectives of PTs who participated in implementing the PT-led primary care model for people with LBP; vi) experiences and perspectives of other primary care team members who participated in implementing the PT-led primary care model for people with LBP; and (vii) how the context influenced implementation of the PT-led primary care model for people with LBP. We plan to present these results at national and international conferences on primary care, health services, physiotherapy, and back pain. We will create tailored summary reports for each manuscript for each of the following knowledge user groups: patients, health professionals, primary care team leaders, and health system decision makers.

Discussion

The results of this trial will inform new models of primary care in Canada and will be applicable to health systems around the world. Our detailed analysis plan has integrated the perspectives of knowledge users, including people living with LBP, primary care providers, PTs, researchers, and health system decision makers. These knowledge users have helped define study outcomes (i.e., disability, quality of life, cost-effectiveness) that will meet the needs of key knowledge users and decision makers. This protocol and analysis plan builds on our trial registration by articulating our analytic decisions in advance of analyzing our data to reduce risk of analytic or reporting bias.

Our knowledge user team has also contributed to our detailed process evaluation, where we aim to explore how the PT-led primary care model for LBP was implemented, potential mechanisms of the model of care, experiences of people involved in implementation, and how context influences implementation. The updated MRC guidance for the development and evaluation of complex interventions emphasizes the importance of moving beyond only questions of effectiveness to also explore how complex interventions, like this model of care, will be accepted, adopted, implemented, scaled and transferred across contexts[64]. Our analysis plan for an in-depth process evaluation aims to support health system decision makers by providing clear and transparent analytic processes to answer the key questions needed to inform scale and spread of this model of care if it is effective.

Acknowledgements: Not applicable.

Funding

This study is funded by a Canadian Institutes of Health Research (CIHR) project grant (reference number 166177). CIHR played no part in the writing of this protocol and analysis plan. Jordan Miller is also supported by an Arthritis Society of Canada and CIHR STARS award which provides financial support for him to focus more of his time on research.

Competing Interest

The authors declare that they have no competing interests.

Data Availability
For access to the data, please contact the corresponding author.

Abbreviations
BC: British Columbia
CFIR: Consolidated Framework for Implementation Research
ED: Emergency Department
EHR: Electronic health record
GEE: Generalized estimating equation
GROC: global rating of change
LBP: Low back pain
MAR: Missing at random
MNAR: Missing not at random
MRC: Medical Research Council (UK)
MSK: Musculoskeletal
PHQ-2: 2-item Patient Health Questionnaire
PSEQ : Pain Self-Efficacy Questionnaire
PT: Physiotherapist
QALY: Quality-adjusted life years
REDCap: Research Electronic Data Capture
REML: Restricted maximum likelihood
RMDQ: Roland-Morris Disability Questionnaire
TSK-11: Tampa Scale of Kinesiophobia 11
WHO: World Health Organization

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. Bone and Joint Canada. Low Back Pain. http://boneandjointcanada.com/low-back-pain/. Published 2014. Accessed August 26, 2022.
3. Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J.* 2008;8(1):8-20.
4. Katz JN. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. *J Bone Joint Surg Am.* 2006;88 Suppl 2:21-24.
5. Gandjour A, Telzerow A, Lauterbach KW, Investigators II. European comparison of costs and quality in the treatment of acute back pain. *Spine (Phila Pa 1976).* 2005;30(8):969-975.
6. Adams NJ, Plane MB, Fleming MF, Mundt MP, Saunders LA, Stauffacher EA. Opioids and the treatment of chronic pain in a primary care sample. *J Pain Symptom Manage.* 2001;22(3):791-796.
7. Reid MC, Engles-Horton LL, Weber MB, Kerns RD, Rogers EL, O'Connor PG. Use of opioid medications for chronic noncancer pain syndromes in primary care. *J Gen Intern Med.* 2002;17(3):173-179.
8. Singer J. *A Snapshot of Health Care in Canada as Demonstrated by Top 10 Lists, 2011.* Canadian Institute of Health Information: Canadian Institute of Health Information;2011.
9. Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. *Spine (Phila Pa 1976).* 2006;31(23):2724-2727.
10. Lim KL, Jacobs P, Klarenbach S. A population-based analysis of healthcare utilization of persons with back disorders: results from the Canadian Community Health Survey 2000-2001. *Spine (Phila Pa 1976).* 2006;31(2):212-218.
11. Jordan KP, Kadam UT, Hayward R, Porcheret M, Young C, Croft P. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. *BMC Musculoskelet Disord.* 2010;11:144.
12. Cote P, Cassidy JD, Carroll L. The treatment of neck and low back pain: who seeks care? who goes where? *Med Care.* 2001;39(9):956-967.
13. Buchbinder R, van Tulder M, Oberg B, et al. Low back pain: a call for action. *Lancet.* 2018;391(10137):2384-2388.
14. Clark S, Horton R. Low back pain: a major global challenge. *Lancet.* 2018;391(10137):2302.
15. Foster NE, Anema JR, Cherkin D, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. *Lancet.* 2018;391(10137):2368-2383.
16. Hartvigsen J, Hancock MJ, Kongsted A, et al. What low back pain is and why we need to pay attention. *Lancet.* 2018;391(10137):2356-2367.
17. Pinney SJ, Regan WD. Educating medical students about musculoskeletal problems. Are community needs reflected in the curricula of Canadian medical schools? *J Bone Joint Surg Am.* 2001;83-A(9):1317-1320.
18. Freedman KB, Bernstein J. Educational deficiencies in musculoskeletal medicine. *J Bone Joint Surg Am.* 2002;84-A(4):604-608.
19. Day CS, Yeh AC, Franko O, Ramirez M, Krupat E. Musculoskeletal medicine: an assessment of the attitudes and knowledge of medical students at Harvard Medical School. *Acad Med.* 2007;82(5):452-457.
20. Darlow B, Dean S, Perry M, Mathieson F, Baxter GD, Dowell A. Acute low back pain management in general practice: uncertainty and conflicting certainties. *Fam Pract.* 2014;31(6):723-732.
21. Petterson SM, Liaw WR, Phillips RL, Jr., Rabin DL, Meyers DS, Bazemore AW. Projecting US primary care physician workforce needs: 2010-2025. *Ann Fam Med.* 2012;10(6):503-509.

22. Paskins Z, Sanders T, Croft PR, Hassell AB. The Identity Crisis of Osteoarthritis in General Practice: A Qualitative Study Using Video-Stimulated Recall. *Ann Fam Med.* 2015;13(6):537-544.
23. Brand CA, Ackerman IN, Tropea J. Chronic disease management: improving care for people with osteoarthritis. *Best Pract Res Clin Rheumatol.* 2014;28(1):119-142.
24. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet.* 2012;380(9836):37-43.
25. Public Health Agency of Canada. *The Chief Public Health Officer's Report on The State of Public Health in Canada 2010 – Growing older: Adding life to years.* 2010.
26. Bath B, Grona SL, Janzen B. A spinal triage programme delivered by physiotherapists in collaboration with orthopaedic surgeons. *Physiother Can.* 2012;64(4):356-366.
27. Dufour SP, Brown J, Deborah Lucy S. Integrating physiotherapists within primary health care teams: perspectives of family physicians and nurse practitioners. *J Interprof Care.* 2014;28(5):460-465.
28. Dufour SP, Lucy SD, Brown JB. Understanding physiotherapists' roles in ontario primary health care teams. *Physiother Can.* 2014;66(3):234-242.
29. Perreault K, Dionne CE, Rossignol M, Morin D. Interprofessional practices of physiotherapists working with adults with low back pain in Quebec's private sector: results of a qualitative study. *BMC Musculoskelet Disord.* 2014;15:160.
30. Saragiotto BT, de Almeida MO, Yamato TP, Maher CG. Multidisciplinary Biopsychosocial Rehabilitation for Nonspecific Chronic Low Back Pain. *Phys Ther.* 2016;96(6):759-763.
31. Delitto A, George SZ, Van Dillen L, et al. Low back pain. *J Orthop Sports Phys Ther.* 2012;42(4):A1-57.
32. Koes BW, van Tulder M, Lin CW, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur Spine J.* 2010;19(12):2075-2094.
33. Pillastrini P, Gardenghi I, Bonetti F, et al. An updated overview of clinical guidelines for chronic low back pain management in primary care. *Joint Bone Spine.* 2012;79(2):176-185.
34. Toward Optimized Practice (TOP) Low Back Pain Working Group. *Evidence-informed primary care management of low back pain: Clinical practice guideline.* 2015 December. Edmonton, AB: Toward Optimizing practice.
35. Comans T, Raymer M, O'Leary S, Smith D, Scuffham P. Cost-effectiveness of a physiotherapist-led service for orthopaedic outpatients. *J Health Serv Res Policy.* 2014;19(4):216-223.
36. James JJ, Stuart RB. Expanded role for the physical therapist. Screening musculoskeletal disorders. *Phys Ther.* 1975;55(2):121-131.
37. Rempel J, Busse JW, Drew B, et al. Patients' Attitudes Toward Non-Physician Screening of Low Back and Low Back-Related Leg Pain Complaints Referred for Surgical Assessment. *Spine (Phila Pa 1976).* 2016.
38. Childs JD, Whitman JM, Sizer PS, Pugia ML, Flynn TW, Delitto A. A description of physical therapists' knowledge in managing musculoskeletal conditions. *BMC Musculoskelet Disord.* 2005;6:32.
39. Hill JC, Dunn KM, Main CJ, Hay EM. Subgrouping low back pain: a comparison of the STarT Back Tool with the Orebro Musculoskeletal Pain Screening Questionnaire. *Eur J Pain.* 2010;14(1):83-89.
40. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet.* 2011;378(9802):1560-1571.
41. Freburger JK, Carey TS, Holmes GM, et al. Exercise prescription for chronic back or neck pain: Who prescribes it? Who gets it? What is prescribed? *Arthritis Care & Research.* 2009;61(2):192-200.
42. Fritz JM, Cleland JA, Brennan GP. Does adherence to the guideline recommendation for active treatments improve the quality of care for patients with acute low back pain delivered by physical therapists? *Med Care.* 2007;45(10):973-980.
43. Rutten GM, Degen S, Hendriks EJ, Braspenning JC, Harting J, Oostendorp RA. Adherence to clinical practice guidelines for low back pain in physical therapy: do patients benefit? *Phys Ther.* 2010;90(8):1111-1122.

44. Riis A, Jensen CE, Bro F, et al. A multifaceted implementation strategy versus passive implementation of low back pain guidelines in general practice: a cluster randomised controlled trial. *Implement Sci.* 2016;11(1):143.
45. Tahvonen P, Oikarinen H, Niinimaki J, Liukkonen E, Mattila S, Tervonen O. Justification and active guideline implementation for spine radiography referrals in primary care. *Acta Radiol.* 2017;58(5):586-592.
46. Fritz JM, Childs JD, Wainner RS, Flynn TW. Primary care referral of patients with low back pain to physical therapy: impact on future health care utilization and costs. *Spine (Phila Pa 1976).* 2012;37(25):2114-2121.
47. Childs JD, Fritz JM, Wu SS, et al. Implications of early and guideline adherent physical therapy for low back pain on utilization and costs. *BMC Health Serv Res.* 2015;15:150.
48. Donato EB, DuVall RE, Godges JJ, Zimmerman GJ, Greathouse DG. Practice analysis: defining the clinical practice of primary contact physical therapy. *J Orthop Sports Phys Ther.* 2004;34(6):284-304.
49. Greathouse DG, Schreck RC, Benson CJ. The United States Army physical therapy experience: evaluation and treatment of patients with neuromusculoskeletal disorders. *J Orthop Sports Phys Ther.* 1994;19(5):261-266.
50. McGill T. Effectiveness of physical therapists servicing as primary care musculoskeletal providers as compared to family practice providers in a deployed combat location: a retrospective medical chart review. *Military Medicine.* 2013;178(10):1115-1120.
51. Murphy BP, Greathouse D, Matsui I. Primary care physical therapy practice models. *J Orthop Sports Phys Ther.* 2005;35(11):699-707.
52. Ross MD, Childs JD, Middel C, et al. Physical therapist vs. family practitioner knowledge of simple low back pain management in the U.S. Air Force. *Mil Med.* 2014;179(2):162-168.
53. Walsh N, Halls S, Thomas R, et al. First Contact Physiotherapy: An evaluation of clinical effectiveness and costs. *Br J Gen Pract.* 2024.
54. Thompson J, Macintosh F, Beaumont N, Bedford L, Powley A, Bailey S. The Experiences and Perceptions of First Contact Practitioners in Primary Care-A Qualitative Systematic Review. *Musculoskeletal Care.* 2024;22(4):e1955.
55. Lamb K, Comer C, Walsh N, Smith J, Tang K, McHugh G. The experiences of patients with musculoskeletal conditions accessing first contact physiotherapy practitioner appointments in general practice in the UK: A qualitative study. *Musculoskeletal Care.* 2024;22(2):e1908.
56. Lamb K, Comer C, Walsh N, McHugh G. Patient access to first contact practitioner physiotherapists in the UK: A national survey. *Musculoskeletal Care.* 2023;21(4):1554-1562.
57. Wood L, Bishop A, Goodwin R, Stynes S. Patient satisfaction with the first contact physiotherapy service: Results from the national evaluation survey. *Musculoskeletal Care.* 2022;20(2):363-370.
58. Goodwin R, Moffatt F, Hendrick P, Stynes S, Bishop A, Logan P. Evaluation of the First Contact Physiotherapy (FCP) model of primary care: a qualitative insight. *Physiotherapy.* 2021;113:209-216.
59. Stynes S, Jordan KP, Hill JC, et al. Evaluation of the First Contact Physiotherapy (FCP) model of primary care: patient characteristics and outcomes. *Physiotherapy.* 2021;113:199-208.
60. Docking S, Sridhar S, Haas R, et al. Models of care for managing non-specific low back pain. *Cochrane Database Syst Rev.* 2025;3(3):CD015083.
61. Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care.* 2003;12(1):47-52.
62. Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ : British Medical Journal.* 2013;346:e7586.
63. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomized trials. *Nat Med.* 2025;31(6):1784-1792.
64. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ.* 2021;374:n2061.

65. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ.* 2015;350:h1258.
66. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. *Implementation Science.* 2022;17(1):75.
67. Kirk MA, Kelley C, Yankey N, Birken SA, Abadie B, Damschroder L. A systematic review of the use of the Consolidated Framework for Implementation Research. *Implement Sci.* 2016;11:72.
68. Breimaier HE, Heckemann B, Halfens RJ, Lohrmann C. The Consolidated Framework for Implementation Research (CFIR): a useful theoretical framework for guiding and evaluating a guideline implementation process in a hospital-based nursing practice. *BMC Nurs.* 2015;14:43.
69. Damschroder LJ, Reardon CM, Opra Widerquist MA, Lowery J. Conceptualizing outcomes for use with the Consolidated Framework for Implementation Research (CFIR): the CFIR Outcomes Addendum. *Implement Sci.* 2022;17(1):7.
70. Vader K, Donnelly C, French SD, et al. Implementing a new physiotherapist-led primary care model for low back pain: a qualitative study of patient and primary care team perspectives. *BMC Prim Care.* 2022;23(1):201.
71. May CR, Mair FS, Dowrick CF, Finch TL. Process evaluation for complex interventions in primary care: understanding trials using the normalization process model. *BMC Fam Pract.* 2007;8:42.
72. Funnell SC, Rogers PJ. *Purposeful program theory: Effective use of theories of change and logic models.* John Wiley & Sons; 2011.
73. Moulton LH. Covariate-based constrained randomization of group-randomized trials. *Clin Trials.* 2004;1(3):297-305.
74. Chaudhary MA, Moulton LH. A SAS macro for constrained randomization of group-randomized designs. *Comput Methods Programs Biomed.* 2006;83(3):205-210.
75. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ.* 2003;327(7418):785-789.
76. Campbell MJ. Challenges of cluster randomized trials. *J Comp Eff Res.* 2014;3(3):271-281.
77. Eldridge S, Kerry S, Torgerson DJ. Bias in identifying and recruiting participants in cluster randomised trials: what can be done? *BMJ.* 2009;339:b4006.
78. Flynn TN, Whitley E, Peters TJ. Recruitment strategies in a cluster randomized trial--cost implications. *Stat Med.* 2002;21(3):397-405.
79. Teerenstra S, Eldridge S, Graff M, de Hoop E, Borm GF. A simple sample size formula for analysis of covariance in cluster randomized trials. *Stat Med.* 2012;31(20):2169-2178.
80. Bombardier C, Hayden J, Beaton DE. Minimal clinically important difference. Low back pain: outcome measures. *J Rheumatol.* 2001;28(2):431-438.
81. Campbell MK, Fayers PM, Grimshaw JM. Determinants of the intracluster correlation coefficient in cluster randomized trials: the case of implementation research. *Clin Trials.* 2005;2(2):99-107.
82. Miller J, Barber D, Donnelly C, et al. Determining the impact of a new physiotherapist-led primary care model for back pain: protocol for a pilot cluster randomized controlled trial. *Trials.* 2017;18(1):526.
83. Thombs BD, Benedetti A, Kloda LA, et al. The diagnostic accuracy of the Patient Health Questionnaire-2 (PHQ-2), Patient Health Questionnaire-8 (PHQ-8), and Patient Health Questionnaire-9 (PHQ-9) for detecting major depression: protocol for a systematic review and individual patient data meta-analyses. *Syst Rev.* 2014;3:124.
84. Linton SJ. Intricacies of good communication in the context of pain: does validation reinforce disclosure? *Pain.* 2015;156(2):199-200.
85. Lumley MA, Sklar ER, Carty JN. Emotional disclosure interventions for chronic pain: from the laboratory to the clinic. *Transl Behav Med.* 2012;2(1):73-81.
86. Gillis ME, Lumley MA, Mosley-Williams A, Leisen JC, Roehrs T. The health effects of at-home written emotional disclosure in fibromyalgia: a randomized trial. *Ann Behav Med.* 2006;32(2):135-146.
87. Pincus T, Holt N, Vogel S, et al. Cognitive and affective reassurance and patient outcomes in primary care: a systematic review. *Pain.* 2013;154(11):2407-2416.

88. Gordon R, Bloxham S. A Systematic Review of the Effects of Exercise and Physical Activity on Non-Specific Chronic Low Back Pain. *Healthcare (Basel).* 2016;4(2).
89. Searle A, Spink M, Ho A, Chuter V. Exercise interventions for the treatment of chronic low back pain: a systematic review and meta-analysis of randomised controlled trials. *Clin Rehabil.* 2015;29(12):1155-1167.
90. Dahm KT, Brurberg KG, Jamtvedt G, Hagen KB. Advice to rest in bed versus advice to stay active for acute low-back pain and sciatica. *Cochrane Database Syst Rev.* 2010(6):CD007612.
91. Main CJ, Sowden G, Hill JC, Watson PJ, Hay EM. Integrating physical and psychological approaches to treatment in low back pain: the development and content of the STarT Back trial's 'high-risk' intervention (StarT Back; ISRCTN 37113406). *Physiotherapy.* 2012;98(2):110-116.
92. Hill JC, Afolabi EK, Lewis M, et al. Does a modified STarT Back Tool predict outcome with a broader group of musculoskeletal patients than back pain? A secondary analysis of cohort data. *BMJ Open.* 2016;6(10):e012445.
93. Louw A, Diener I, Butler DS, Puentedura EJ. The effect of neuroscience education on pain, disability, anxiety, and stress in chronic musculoskeletal pain. *Arch Phys Med Rehabil.* 2011;92(12):2041-2056.
94. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform.* 2019;95:103208.
95. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-381.
96. Groll DL, Heyland DK, Caeser M, Wright JG. Assessment of Long-Term Physical Function in Acute Respiratory Distress Syndrome (ARDS) Patients: Comparison of the Charlson Comorbidity Index and the Functional Comorbidity Index. *American Journal of Physical Medicine & Rehabilitation.* 2006;85(7):574-581.
97. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. *J Clin Epidemiol.* 2005;58(6):595-602.
98. Rundell SD, Resnik L, Heagerty PJ, Kumar A, Jarvik JG. Performance of the Functional Comorbidity Index (FCI) in Prognostic Models for Risk Adjustment in Patients With Back Pain. *PM R.* 2020;12(9):891-898.
99. Chiarotto A, Maxwell LJ, Terwee CB, Wells GA, Tugwell P, Ostelo RW. Roland-Morris Disability Questionnaire and Oswestry Disability Index: Which Has Better Measurement Properties for Measuring Physical Functioning in Nonspecific Low Back Pain? Systematic Review and Meta-Analysis. *Phys Ther.* 2016.
100. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine (Phila Pa 1976).* 2000;25(24):3115-3124.
101. Haefeli M, Elfering A. Pain assessment. *Eur Spine J.* 2006;15 Suppl 1(Suppl 1):S17-S24.
102. van Reenen M, Janssen B, Stolk E, et al. EuroQol Research Foundation. EQ-5D-5L User Guide, 2019. doi:https://euroqol.org/publications/user-guides.
103. Finch AP, Dritsaki M, Jommi C. Generic Preference-based Measures for Low Back Pain: Which of Them Should Be Used? *Spine (Phila Pa 1976).* 2016;41(6):E364-374.
104. Xie F, Pullenayegum E, Gaebel K, et al. A Time Trade-off-derived Value Set of the EQ-5D-5L for Canada. *Med Care.* 2016;54(1):98-105.
105. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther.* 2009;17(3):163-170.
106. Fritz JM, Hebert J, Koppenhaver S, Parent E. Beyond minimally important change: defining a successful outcome of physical therapy for patients with low back pain. *Spine (Phila Pa 1976).* 2009;34(25):2803-2809.
107. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Pilot Feasibility Stud.* 2016;2:64.
108. Ioannidis JP, Evans SJ, Gotzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med.* 2004;141(10):781-788.

109. Chiarotto A, Vanti C, Cedraschi C, et al. Responsiveness and Minimal Important Change of the Pain Self-Efficacy Questionnaire and Short Forms in Patients With Chronic Low Back Pain. *J Pain.* 2016;17(6):707-718.
110. Nicholas MK. The pain self-efficacy questionnaire: Taking pain into account. *Eur J Pain.* 2007;11(2):153-163.
111. Osman A, Barrios FX, Kopper BA, Hauptmann W, Jones J, O'Neill E. Factor structure, reliability, and validity of the Pain Catastrophizing Scale. *J Behav Med.* 1997;20(6):589-605.
112. Sullivan MJ, Thorn B, Haythornthwaite JA, et al. Theoretical perspectives on the relation between catastrophizing and pain. *Clin J Pain.* 2001;17(1):52-64.
113. Roelofs J, Goubert L, Peters ML, Vlaeyen JW, Crombez G. The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. *Eur J Pain.* 2004;8(5):495-502.
114. Swinkels-Meewisse EJ, Swinkels RA, Verbeek AL, Vlaeyen JW, Oostendorp RA. Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. *Man Ther.* 2003;8(1):29-36.
115. Manea L, Gilbody S, McMillan D. A diagnostic meta-analysis of the Patient Health Questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression. *Gen Hosp Psychiatry.* 2015;37(1):67-75.
116. Ministry of Health. OHIP Schedule of Benefits and fees. https://www.ontario.ca/page/ohip-schedule-benefits-and-fees. Published 2024. Updated August 30, 2024. Accessed November 13, 2024.
117. Drummond M, Scupher M, Claxton K, Stoddard G, Torrance G. *Methods for the economic evaluation of health care programs.* Oxford, UK: Oxford University Press.
118. Gomes M, Grieve R, Nixon R, Edmunds WJ. Statistical methods for cost-effectiveness analyses that use data from cluster randomized trials: a systematic review and checklist for critical appraisal. *Med Decis Making.* 2012;32(1):209-220.
119. Gomes M, Ng ES, Grieve R, Nixon R, Carpenter J, Thompson SG. Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials. *Med Decis Making.* 2012;32(2):350-361.
120. Grieve R, Nixon R, Thompson SG. Bayesian hierarchical models for cost-effectiveness analyses that use data from cluster randomized trials. *Med Decis Making.* 2010;30(2):163-175.
121. National Guideline Centre NICE. *Low Back Pain and Sciatica in Over 16s: Assessment and Management.* London: National Institute for Health and Care Excellence (NICE);2016.
122. Grant A, Kontak J, Jeffers E, et al. Barriers and enablers to implementing interprofessional primary care teams: a narrative review of the literature using the consolidated framework for implementation research. *BMC Prim Care.* 2024;25(1):25.
123. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies: Guided by Information Power. *Qual Health Res.* 2016;26(13):1753-1760.
124. Kenward MG, Roger JH. An improved approximation to the precision of fixed effects from restricted maximum likelihood. *Computational Statistics & Data Analysis.* 2009;53(7):2583-2595.
125. Hooper R, Forbes A, Hemming K, Takeda A, Beresford L. Analysis of cluster randomised trials with an assessment of outcome at baseline. *BMJ.* 2018;360:k1121.
126. Hemming K, Taljaard M. Key considerations for designing, conducting and analysing a cluster randomized trial. *International Journal of Epidemiology.* 2023;52(5):1648-1658.
127. Shmagel A, Foley R, Ibrahim H. Epidemiology of Chronic Low Back Pain in US Adults: Data From the 2009-2010 National Health and Nutrition Examination Survey. *Arthritis Care Res (Hoboken).* 2016;68(11):1688-1694.
128. George SZ, Fritz JM, Childs JD, Brennan GP. Sex differences in predictors of outcome in selected physical therapy interventions for acute low back pain. *J Orthop Sports Phys Ther.* 2006;36(6):354-363.
129. Mathieu J, Roy K, Robert ME, Akeblersane M, Descarreaux M, Marchand AA. Sociodemographic determinants of health inequities in low back pain: a narrative review. *Front Public Health.* 2024;12:1392074.

130. Bekkering GE, Hendriks HJ, van Tulder MW, et al. Prognostic factors for low back pain in patients referred for physiotherapy: comparing outcomes and varying modeling techniques. *Spine (Phila Pa 1976).* 2005;30(16):1881-1886.
131. Moreno-Betancur M, Chavance M. Sensitivity analysis of incomplete longitudinal data departing from the missing at random assumption: Methodology and application in a clinical trial with drop-outs. *Statistical Methods in Medical Research.* 2016;25(4):1471-1489.
132. Troendle JF, Sur A, Leifer ES, Powell-Wiley T. Sensitivity Analyses for Missing in Repeatedly Measured Outcome Data. *Stat Med.* 2025;44(23-24):e70282.
133. Chavance M. Handling Missing Items in Quality of Life Studies. *Communications in statistics Theory and methods.* 2004;33(6):1371-1383.
134. Yelland LN, Salter AB, Ryan P. Performance of the Modified Poisson Regression Approach for Estimating Relative Risks From Clustered Prospective Data. *American Journal of Epidemiology.* 2011;174(8):984-992.
135. Fay MP, Graubard BI. Small-Sample Adjustments for Wald-Type Tests Using Sandwich Estimators. *Biometrics.* 2004;57(4):1198-1206.
136. Henschke N, van Enst A, Froud R, Ostelo RW. Responder analyses in randomised controlled trials for chronic low back pain: an overview of currently used methods. *Eur Spine J.* 2014;23(4):772-778.
137. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain.* 2008;9(2):105-121.
138. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976).* 2008;33(1):90-94.
139. Qiu H, Cook AJ, Bobb JF. Evaluating tests for cluster-randomized trials with few clusters under generalized linear mixed models with covariate adjustment: A simulation study. *Stat Med.* 2024;43(2):201-215.
140. Li P, Redden DT. Small sample performance of bias-corrected sandwich estimators for cluster-randomized trials with binary outcomes. *Statistics in medicine.* 2015;34(2):281-296.
141. Thompson JA, Leyrat C, Fielding KL, Hayes RJ. Cluster randomised trials with a binary outcome and a small number of clusters: comparison of individual and cluster level analysis method. *BMC Medical Research Methodology.* 2022;22(1):222.
142. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. *Res Integr Peer Rev.* 2016;1:2.
143. O'Sullivan K. Appraisal of Clinical Practice Guideline: World Health Organization guideline for non-surgical management of chronic primary low back pain in adults in primary and community care settings. *Journal of Physiotherapy.* 2024;70(2):156.
144. Buchbinder R, Underwood M, Hartvigsen J, Maher CG. The Lancet Series call to action to reduce low value care for low back pain: an update. *Pain.* 2020;161 Suppl 1(1):S57-S64.
145. Hartvigsen J, Kamper SJ, French SD. Low-value care in musculoskeletal health care: Is there a way forward? *Pain Pract.* 2022;22 Suppl 2(Suppl 2):65-70.
146. Hayes AF. *Introduction to mediation, moderation, and conditional process analysis: A regression based approach.* Second ed: Guilford Press; 2017.
147. Berli C, Inauen J, Stadler G, Scholz U, Shrout PE. Understanding Between-Person Interventions With Time-Intensive Longitudinal Outcome Data: Longitudinal Mediation Analyses. *Ann Behav Med.* 2021;55(5):476-488.
148. VanderWeele TJ. A three-way decomposition of a total effect into direct, indirect, and interactive effects. *Epidemiology (Cambridge, Mass).* 2013;24(2):224-232.
149. VanderWeele TJ. Mediation Analysis: A Practitioner's Guide. *Annu Rev Public Health.* 2016;37:17-32.
150. Zhao X, Lynch JG, Jr., Chen Q. Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research.* 2010;37(2):197-206.

151. Preacher KJ. Advances in Mediation Analysis: A Survey and Synthesis of New Developments. *Annual Review of Psychology.* 2015;66(1):825-852.
152. Selig JP, Preacher KJ. Mediation Models for Longitudinal Data in Developmental Research. *Research in Human Development.* 2009;6(2-3):144-164.
153. Goldsmith KA, MacKinnon DP, Chalder T, White PD, Sharpe M, Pickles A. Tutorial: The practical application of longitudinal structural equation mediation models in clinical trials. *Psychological Methods.* 2018;23(2):191-207.
154. Ananth CV. Proportion mediated in a causal mediation analysis: how useful is this measure? *BJOG.* 2019;126(8):983.
155. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. *International Journal of Epidemiology.* 2013;42(5):1511-1519.
156. VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology.* 2010;21(4):540-551.
157. Hunt MR. Strengths and challenges in the use of interpretive description: reflections arising from a study of the moral experience of health professionals in humanitarian work. *Qual Health Res.* 2009;19(9):1284-1292.
158. Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a noncategorical qualitative alternative for developing nursing knowledge. *Res Nurs Health.* 1997;20(2):169-177.
159. Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. *International journal of qualitative methods.* 2002;1(2):13-22.
160. Holloway I, Galvin K. *Qualitative research in nursing and healthcare.* John Wiley & Sons; 2016.
161. Guba EG, Lincoln YS. *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches.* Jossey-Bass; 1981.
162. Lincoln YS, Guba EG. But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New directions for program evaluation.* 1986;1986(30):73-84.
163. Guba EG. Criteria for assessing the trustworthiness of naturalistic inquiries. *Ectj.* 1981;29(2):75.
164. Reardon CM, Damschroder LJ, Ashcraft LE, et al. The Consolidated Framework for Implementation Research (CFIR) User Guide: a five-step guide for conducting implementation research using the framework. *Implement Sci.* 2025;20(1):39.
165. Damschroder LJ, Goodrich DE, Robinson CH, Fletcher CE, Lowery JC. A systematic exploration of differences in contextual factors related to implementing the MOVE! weight management program in VA: a mixed methods study. *BMC Health Serv Res.* 2011;11:248.
166. Damschroder LJ, Reardon CM, AuYoung M, et al. Implementation findings from a hybrid III implementation-effectiveness trial of the Diabetes Prevention Program (DPP) in the Veterans Health Administration (VHA). *Implement Sci.* 2017;12(1):94.
167. Damschroder LJ, Reardon CM, Sperber N, Robinson CH, Fickel JJ, Oddone EZ. Implementation evaluation of the Telephone Lifestyle Coaching (TLC) program: organizational factors associated with successful implementation. *Transl Behav Med.* 2017;7(2):233-241.
168. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. *Health Serv Res.* 2013;48(6 Pt 2):2134-2156.