testing (e.g. advanced imaging), and limit exposure to higher risk treatment pathways (e.g. opioids, injections, and/or surgery).[13]

Table 2. Distinctions of the Two Care Pathways

	Integrated Care	**Coordinated Care**
Initiation	Provider referral	Provider referral
Staff Contact	Physical Therapist	Pain Navigator
Provider Approach	Directed care - Pain modulation Home based activity, Risk stratification, Behavioral/psychologically informed treatment (if indicated)	Coordinated care - Patient preference Guideline adherence Facilitated referrals to existing VA or non-VA pain management resources
Care Progression	Sequenced	Stepped

2.2 Study Rationale

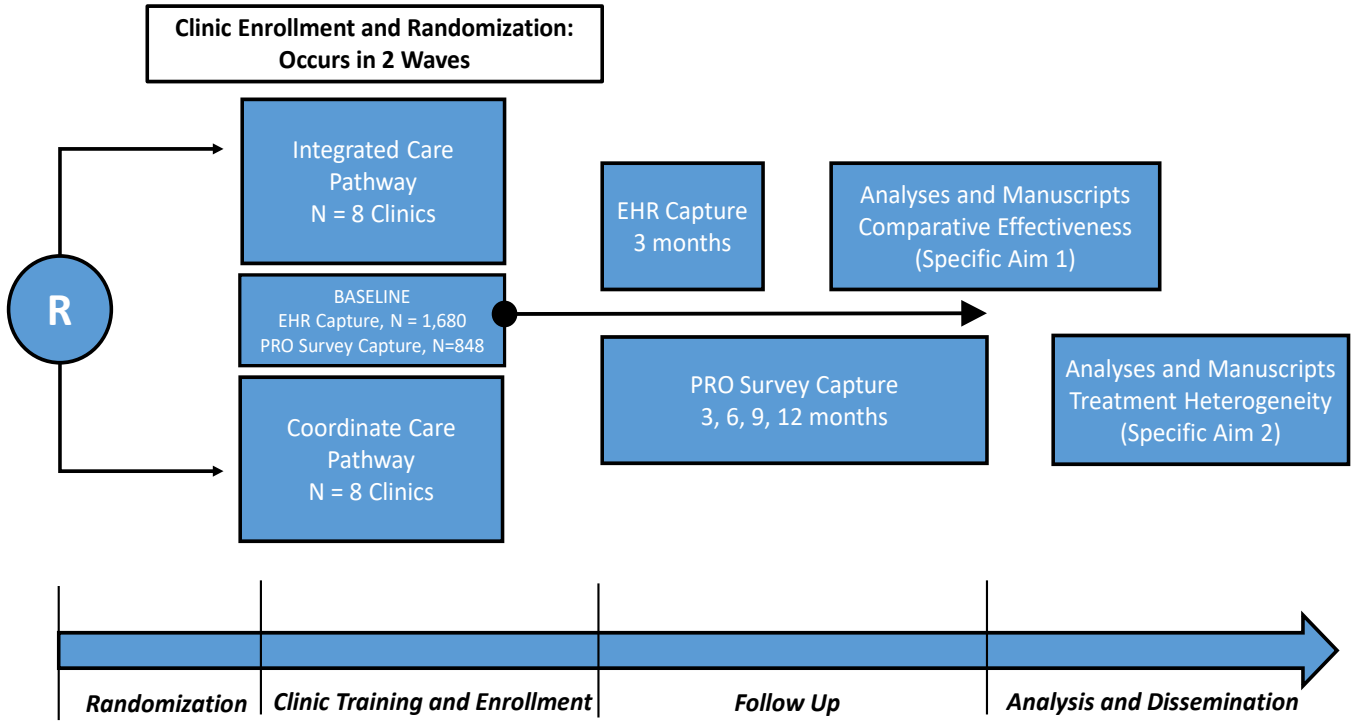
Among conditions targeted for improved non-pharmacological pain management strategies, musculoskeletal pain, and LBP specifically, is a particular priority.[5,6] Of those affected, Veterans are more likely to have LBP or joint pain, and are more likely to report LBP or joint pain as severe when experienced.[8] The widespread impact of LBP and the disproportionate impact of LBP on Veterans' quality of life argue strongly for improvement in non-pharmacological pain management care pathways in the VA setting. Indeed, new care pathways that meet the needs of Veterans with LBP and are associated with favorable clinical outcomes represent a critical step forward in patient care.[13] To advance understanding relevant to the above priorities, we will conduct a cluster randomized embedded pragmatic trial to develop infrastructure and investigate the comparative effectiveness of two different pain management strategies that have potential to improve access to non-pharmacological care for LBP. The two pain management care pathways that will be evaluated in this trial are 1) an integrated care pathway (ICP) and 2) a coordinated care pathway (CCP). These two strategies share the common goal of leveraging existing VA infrastructure for non-pharmacological pain management of low back pain, but the structure of care delivery is notably different in terms of ***staff contact, provider approach, and care progression*** (**Table 2**).

In this trial, we will determine whether these differences impact patient outcomes and/or health care utilization. Evaluating two pain care pathways in a well-powered comparative effectiveness trial will help guide infrastructure development and contribute to novel understanding of implementation and performance in different clinical settings and situations. Other significant aspects of the proposed pain management strategies, including structuring care and following recommendations for biopsychosocial care and reducing variation in care, are further discussed below.

In recognition of the high burden of pain on Veterans and associated costs to the health care system, VHA implemented a National Pain Management Strategy.[1] This Strategy approaches pain care within a biopsychosocial framework and this has led to VA offering a broad array of services for pain care. For example, VHA has specifically encouraged expansion of access to cognitive behavioral therapy for chronic pain management. However, a major barrier to organized and timely delivery of these services has been limited provider and Veteran awareness of them. Other barriers include lack of a standardized process for connecting patients to these services (e.g., who is responsible, which services are appropriate). In addition, recent legislation affecting financing of services outside of VA (e.g. CHOICE Act) has created a rapidly changing environment regarding how these services are organized, delivered, and paid.

VA primary care clinics that have agreed to have their clinical staff trained in offering these care pathways as standard care will be randomly assigned to one of the two care pathways. A cluster randomized trial (CRT) was selected because it offers the most pragmatic and efficient design to address our aims due to the nature of the care pathways and interest in assessing both short and long-term outcomes. An important consideration in a CRT is the optimal unit of randomization. We will designate the unit of randomization as the VA clinic (rather than an entire VAMC 'site' or 'station') to 1) enhance feasibility of implementation, and 2) improve power (higher number of clusters).

Figure 3. Study Design Overview



Veteran outcomes will be collected in the EHR at baseline and at 3-months follow up and for a subset of Veterans, surveys will be administered at baseline with follow up points at 3, 6, 9, and 12 months (see Section 10 - Measures and Data Collection). The primary outcomes (PROMIS short forms for pain interference and physical function, 4 items each) and secondary outcome (PROMIS short form sleep quality) will be captured in the VA medical record by clinical providers using templates provided by the research team (see Section 6 - Study Procedures). Veteran participants (n=848) will be consented for the additional survey data collection by Durham VA study staff (baseline) and DCRI's Outcomes Call Center (3, 6, 9, and 12 month follow-up calls).

This trial is purposefully powered to examine treatment effect moderation for care strategy responses in select patient subgroups. Many clinical trials find negligible or small treatment effects for LBP interventions, so identification of treatment moderators (i.e. factors that predict largest benefit from intervention) is important for a better understanding of how treatment could be administered at the patient level. This is a newer area of emphasis in LBP research and there are methodological[38] and statistical[39] challenges to investigating treatment moderation. We have mitigated many of those challenges by taking an *a priori* approach that

1. Currently in institutional care (nursing home or hospital).
2. Cognitive impairment or dementia (identified via ICD diagnosis codes or PCP note in previous 2 years) or lack decision-making capacity, documented in the medical record.
3. Serious mental illness defined as diagnosis of schizophrenia, bipolar disorder, psychiatric hospitalization in the previous year or current high-risk suicide flag in their CPRS medical record.
4. Unable to communicate on the telephone, or no telephone access for duration of study

Opt-out letter and enrollment over the telephone. At the beginning of the study, centrally located Durham VA study staff will generate an opt-out letter signed by the research team that will be mailed to all potentially eligible patients enrolled in participating VA primary care clinics. The letter will describe the study and include an "opt-out" phone number if they do not wish to be contacted further. We will also include opt-out information on printed clinical program brochures that providers give patients in clinic at the time of referral.

We have opted not to enroll Veterans in person because phone interviews improve feasibility given the number of (n = 16 clinics), their expected geographic dispersion, and the overall number of participants that will need to provide consent for survey contact (n = 848). In addition, enrolling by phone is least burdensome for patients and allows us to reach patients whose clinic visit may have occurred on a weekend or during evening hours. All VAMC primary care clinics now offer some expanded hours.

Screening, Informed Consent, Baseline Data Collection. Durham VA study staff will conduct a biweekly data pull that identifies Veterans that have been referred to the pathway at enrolled clinics who meet eligibility criteria. We will randomly sort eligible participants, stratified by clinic, and the RAs will recruit from this list for the Telephone Survey Subset. Before contacting each patient for enrollment, RAs will perform targeted chart review to confirm eligibility and then contact Veterans to screen for additional exclusion criteria not always identified in the medical record (e.g. hearing impairment or hospice referral) and ask them if they wish to participate. Because the study is entirely telephone based, waiver of written informed consent has been granted by the Durham VAMC IRB. After obtaining verbal informed consent, the Durham VA RA will collect baseline survey data from all study participants. Study staff will attempt to contact Veterans for enrollment within 2-3 days of the data pull date and make up to 10 attempts to contact each potentially eligible subject.

Telephone Surveys Follow-up.

Survey participants will have provided verbal informed consent prior to baseline survey collection to allow for additional contact from the Duke Outcomes Call Center for the follow-up surveys at 3,6,9 and 12 months.

Data collected via telephone interviews will complement the data obtained via EHR with patient-reported data across multiple domains of pain and associated co-morbidities. These data will also provide rich information about response trajectories and suggest areas to target future EHR data capture methods. This approach ensures the trial will build capacity for conducting large-scale pragmatic clinic trials with the VA health care system.

7.2 Safety Monitoring

Because this is collecting standard of care information, and contact with the participants only involves telephone survey-based assessments, this is considered a minimal risk study. There is no investigational drug or device, and all data collected through EMR records are considered standard of care data.

The DSMB will be implemented to review our research procedures and operating procedures before initiating the study and to ensure the data collection and safety-monitoring are implanted while the study is ongoing. The DSMB will also be involved with reporting of any adverse events that may occur, and will make the determination of whether study stoppage is indicated. Currently we do not have any interim analyses planned for this proposal, but we will pursue this option if it is indicated by the DSMB. The DSMB will meet at least annually either in person or on teleconference, and will meet more frequently if circumstances warrant. For example, in the situation when multiple adverse events may merit a reconsideration of the risk to benefit ratio for this study. The DSMB will likely consist of 3-5 individuals with the appropriate expertise to review this study, but who are not directly involved with its operation and are not considered as prior collaborators with this research team. The minutes and executive summary of each DSMB meeting will be submitted to the NCCIH, the NIH Pain Management Work Group and the local IRB's involved with this trial.

8. CLINIC DISCONTINUATION

While our goal is to not have any clinics discontinue, clinic discontinuation could occur by a clinic informing us that they were dropping out after randomization but before training and rollout of pathways or after rollout of pathways or some other unplanned circumstance. It is important to note that our design allows for additional sites to be added in the 2nd round of cluster randomization if site(s) from the first round discontinue or have lower than expected recruitment rates. This is described in more detail in Section 9 (Statistical Considerations).

One of the eligibility criteria for clinics is that the volume of back patients seen on a weekly basis is on average 8 -12 so that over a 6-month period assuming a 50% referral rate (4-6 referred per week) clinics can meet recruitment goal of 105 Veterans referred to their pathway; it may take clinics with smaller patient volumes or lower referral rates longer to enroll and weekly referrals could vary from 2 to 4 per week and we can still meet recruitment goals on our timeline.

Given that this is a embedded pragmatic trial and "enrollment" is controlled by referrals from primary care providers to the clinics randomized pathway either CCP or ICP, the study team will monitor care pathway volume monthly at each participating clinical site. The goal to monitoring pathway volume will be for the study team to trouble-shoot problems to help clinics increase pathway volume. Over the first couple of quarters of an enrollment period at a clinic we will review number of referrals with clinics to make sure meeting goals and to trouble-shoot any problems or issues. If it appears that lower than expected volume is occurring mitigating steps will include meeting with site leaders to review care pathway procedures, trouble-shoot, and enhance education outreach for referring providers. At the time that we begin enrolling block 2 clinics, we will add additional clinics if any has dropped out in block 1 and we will examine recruitment in the clinics in block 1 and determine whether we need to enroll and randomize additional clinics in block 2 to compensate for low recruitment (below target of 105 per clinic) at any of the block 1 clinics. The blocked covariate constrained randomization balances covariates within and across blocks. We will use any information from the enrollment process of the 10 clinics in block 1 to inform and improve the enrollment process for the set of clinics for block 2.

Study measures were selected to be patient-centered and consistent with recent recommendations from the NIH Task Force on Research Standards for Chronic Back Pain, and relevant to Veteran's health and VHA.[49]

9.2 Sample Size and Randomization

In **Table 8** below we present a range of minimum detectable effect size differences for the primary outcomes of Pain Interference or Function Scores for both the administrative sample (N=84 per clinic) and survey sample (N=42 per clinic) for 16 clinics (8 clinics randomized to each arm) for Aim 1. Sample size calculations for the cluster randomized design were based on the net difference between the two conditions across baseline and 3-months follow-up.[50] We assumed over time correlations of 0.50 for both patients (based on data from the Goode et al. pilot study[33] and summary data available for Durham VA clinics, intraclass correlation coefficients (ICC) of 0.01, 0.02, and 0.05 to account for clinic clustering, and also adjusted for the potential group-level randomization of up to 4 stratification variables as baseline covariate adjustments where we conservatively assumed no reduction in either group or subject level variance components. For all calculations, the type-I error is 2.5% to account for the multiple primary outcomes and power is conservatively assumed to be 90% to guard against deviations from assumptions. The number needed at baseline (N_b) is based on an attrition rate of 20%. Note in the **below Table 4** that the scenarios above lead to having adequate power to detect a medium effect size for the primary outcomes for both the administrative and survey samples.

Table 8. Minimum detectable effect size differences for 90% power, alpha of 0.025 and ICC of 0.01, 0.02, and 0.05 for administrative outcome (n=84 per clinic) and survey outcome (n=42 per clinic)

Number of clinics per condition	Number of patients per clinic needed at post-intervention follow-up (Total N)	Minimum Detectable Effect Size Difference ICC=0.01	Minimum Detectable Effect Size Difference ICC=0.02	Minimum Detectable Effect Size Difference ICC=0.05	Number of patients per clinic needed at baseline with 20% attrition (Total N_b)
8	42 (N=672)	0.37	0.42	0.55	53 (N_b=848)
8	84 (N=1344)	0.30	0.36	0.50	105 (N_b=1680)

The standardized effect sizes[51] that we are powered to detect range from 0.30 to 0.50 for the primary outcomes on the administrative sample. This range maps to differences of approximately 2.4 to 5.0 points in the PROMIS Pain Interference score between arms assuming standard deviations in the range found in the pilot study from Goode et al.[33] These magnitude of differences have been reported to be clinically relevant.[52] Similarly for the survey sample, the range of effect size differences we can detect map to differences of approximately 3.0 to 5.5 points between arms.

For AIM 2, the subgroup analyses of the primary outcomes in the administrative sample, if we assume an even distribution of a binary moderator (e.g. opioid use), our effective sample size for detection of moderation is a quarter that for the main effect analysis yielding an effective sample size of n=336 (21 per clinic).[53] For the moderator interaction effect, we can detect medium to large effect size differences over a range of assumptions (**Table 9 below**), which allows for prudent exploration of treatment moderation.

9.5.2 Secondary Outcomes

Secondary outcomes of PROMIS sleep, opioid use, and utilization will also be obtained administratively on the full administrative sample (n=1680). As previously describe for the primary measure PROMIS sleep will be collected at the same times as the primary outcomes in the EHR (see Section 6). Opioid use will be defined at baseline and then at 12 months later using the PMC3 provided definition. For imaging and ED utilization, we will measure use for the 12 months following referral.

9.6 Data Analyses

Aim 1 Analyses

The primary outcomes are continuous and will be ascertained at the planned baseline and follow-up assessment (3 months) from administrative data collected from the EHR on all eligible LBP Veterans presenting at participating clinics. Changes in pain interference and/or physical function scores will be estimated and the primary hypotheses tested via hierarchical linear mixed-effects models with patients nested within clinics and baseline and 3 month values in the response vector.[55] Hierarchical linear models are a flexible and powerful analytic tool for clustered longitudinal continuous outcomes. The fixed-effect portion of the model will have the form: $Y_{ijk} = \beta_0 + \beta_1*(followup) + \beta_2*(followup*intervention)$ for clinic *i*, patient *j,* at time *k*. Random effects (clinics and time by clinics) will be included in the model to account for clustering of patients within clinics as the clinics are the unit of randomization, random effects will also be included to account for the within-patient correlation between repeated measures over time. We will assess the best fitting random effects structure by fitting a variety of random coefficient models (e.g., random intercept only, random intercept and linear slope) and assess using Akaike Information Criteria (AIC) model selection criteria.[56,57] The predictors in the model will include a time effect and indicator variables for treatment interacting with the time effect. The intraclass correlation capturing the relationship of outcomes between patients seen at the same clinic is accounted for via the random effects for the clinics and time by clinics, which are assumed to be normally distributed. The model will be fit in the SAS procedure PROC MIXED using full likelihood approximation and the hypotheses will be tested by whether the estimated coefficient β_2 is positive and significantly different than 0 at the 0.025 level due to 2 primary outcome variables. We will include covariates used in the covariate constrained randomization[4] (5 potential variables; average pain scores, clinic location (main medical center/community clinic), number of participating primary care providers, average level of opioid exposure of LBP patients at clinic, and average age of LBP patients at clinic) in our primary model as well as a limited number of patient-level covariates that are readily available in the EHR (age, gender, race and a comorbidity measure). While we have procedures in place to monitor timing of the 3-month follow-up assessment in the EHR, if we have a significant number of assessments outside the 1-month window for the 3-month assessment that are dropped from the primary analysis, we will conduct a sensitivity analysis including the assessments outside the window and estimating a 3-month treatment effect from this model.

Secondary Analyses

We will assess differences in proportions of “responders” between treatment arms for each of the primary outcomes. A “responder” will be defined as achieving 30% improvement in pain interference or function scores at 3 months follow up.[58] Responder status for those missing outcome data will be estimated using best linear unbiased predictors from the hierarchical linear model [59]so that those missing observed follow up data will be included.

We will use a generalized linear mixed model with a logit link function to compare differences in responder status between arms where the main predictor of interest will be treatment arm adjusting for clustering of VA clinic either with random effects or by conditioning.[55,60,61] As a sensitivity analysis to explore responder cut points of pain interference and function measures, we will compute cumulative proportion responder analysis graphs.[62]

Secondary outcomes of sleep, opioid use, and health utilization (including chronic opioid usage) will also be obtained administratively on all eligible subjects. The sleep PROMIS measure is a continuous outcome that will be assessed at baseline and 3 months and similar modeling procedures as described for the primary outcomes will be used. Opioid use will be examined in two ways, one as a binary variable based on whether a chronic opioid user or not at baseline and 12 months and second using a continuous measure of morphine equivalents for opioid dose at baseline and 12 months.[63] For the dichotomous outcomes, we will use a generalized linear mixed model with a logit link function where the main predictor of interest will be treatment arm and will include baseline opioid use status (chronic or not chronic) adjusting for clustering of VA clinic either with random effects or by conditioning.[55,60,61] For opioid morphine equivalent dose, we will fit a similar model as was described for the primary outcomes except the follow up time point in the model will be for 12 months. For imaging and ED utilization, we will assume a count-like distribution for the number of ED visits in the 12 months of follow up and will use generalized linear mixed model for count variables.[55]

The survey outcomes for the enrolled subset of patients will be collected at baseline, 3, 6, 9, and 12 months and will include pain interference, function, intensity, catastrophizing, sleep, and depression (see Section 10 for full descriptions). These are all longitudinal continuous outcomes and a hierarchical linear model similar to that described for the primary aim will be fit. We will fit random coefficient models as described above (e.g. random intercept only, random intercept and linear slope) and assess using AIC model selection criteria to determine best model for the covariance structure. Similarly, we will determine the best model for the mean structure (e.g. linear, quadratic, dummy coding) as there are five outcome measurement occasions guided by descriptive plots and model fit assessed using AIC model selection criteria. Due to the timing of administration of baseline surveys that some baseline surveys may occur after the initial intervention contact, we will conduct sensitivity analysis treating baseline surveys after initial provider contact as occurring in the post-treatment period.

Aim 2 Analyses

As highlighted in the UG3 section we have strong clinical and scientific rationale to investigate chronic vs. acute LBP and previous opioid use vs. opioid naïve as *a priori* subgroups for treatment moderation.[38] We will define "chronic low back" as pain that has persisted for at least 3 months and resulted in pain on at least half the days in the past six months to be consistent with the NIH task force definition.[49] All participants will have chronicity of back pain assessed via the EHR template (see Section 6) at the time of initial referral to a given care pathway using questions specified in the NIH minimum dataset.[49] Opioid exposure prior to pathway entry will be defined as being prescribed at least one prescription opioid for 20 or more consecutive days (yes or no within 12 months prior to entering pathway).[64] This definition could be altered if different ones are suggested for use by the PMC3. These data will be obtained through VA's Pharmacy Benefits Management System.

For the planned subgroup analyses for key moderators on the primary outcomes of pain interference and function will be the same modeling framework as described for AIM 1 with the addition of the indicator variable(s) for the moderator and associated interactions.[65] The fixed-effect portion of the model will have the form: $Y_{ijk} = \beta_0 +$

	PROMIS SF - 4 item	Patient report on CATI survey	x	x	x	x	X
Sedative/hypnotic exposure	At least one prescription sedative/hypnotic for 20 or more consecutive days	Pharmacy Benefits Management	x				X
PTSD symptoms	PCL-5	CDW health factor data element	x		X		X
Depressed Mood	PHQ-2 Questionnaire	Patient report on CATI survey	x		x		X
Alcohol	AUDIT-C	Patient report on CATI survey	x				
Utilization Measures							
Imaging, Providers, and Procedures	Utilization - Radiographs, MRI, CT - Provider types and number - Injection, ablation, and surgery	CDW inpatient & outpatient Non-VA care (Fee) CPT outpatient + ICD codes					x
ED visits	ED visits - VA - Non-VA	CDW inpatient & outpatient Non-VA care (Fee) inpatient & outpatient					X
Pain-related ED visits	ED visits - VA - Non-VA	CDW inpatient & outpatient files; Non-VA care (Fee) inpatient & outpatient + ICD codes					X

10.1 Survey Data Collection Process

Because the participation in the two care pathways described earlier in the protocol is considered a quality improvement initiative (see memo in appendix), local site activities and participation in the two care pathways will not require IRB oversight. Two separate (Durham VA and Duke University) IRBs have been approved for described research related project activities. The Durham VAMC IRB will be the IRB of record to cover participant survey recruitment activities, including obtaining verbal informed consent for participating in the telephone survey, administering a telephone survey for baseline data collection activities, medical record chart abstraction, and post data collection analysis activities. The Duke University IRB will cover follow up data collection activities performed by the Duke Clinical Research Institute Call Center.

As described earlier in Section 4, centrally located Durham VA study staff will generate an opt-out letter signed by the research team that will be mailed to all potentially eligible patients enrolled in participating VA primary care clinics and a query will be run biweekly by the Durham VA study staff team to identify eligible patients based on inclusion and exclusion criteria. Patients who have not opted-out will be contacted by Durham VA study staff to consent and collect baseline data and verbally consent to be interviewed every three months by the Duke Clinical Research Institute call center staff. The table above documents the potential questions which will be asked in follow up interviews. These questions will serve as the CRFs for the follow up data collection activities.

Pre-consent screening data will be collected in the VA instance of REDCap survey software. Following verbal consent, baseline assessments will also be collected by VA staff using the VA instance of REDCap. Post-consent, follow-up assessments will be conducted by the DCRI Call Center. Follow-up survey data will be collected and stored

This has increased confusion about how and when pain-related services should be offered to Veterans.

VA clinics that have agreed to have their clinical staff trained in offering these care pathways as standard care will be randomly assigned to one of the two care pathways. A cluster randomized trial (CRT) was selected because it offers the most pragmatic and efficient design to address our aims. VA clinics implementing these back pain care pathways are not engaged in research because both pathways involve different ways of providing standard of care treatment. Employees at clinical sites that implement the back pain care pathways will participate in educational training for delivering the treatment as part of standard care and clinical program-related duties only. Employees delivering the clinical treatment will not be directly involved in any research related procedures (e.g. extraction of data from VA's Corporate Data Warehouse or telephone surveys). To assist each site in implementing the care pathways we will collect data and provide reports to be used for process improvement. These will include both quantitative (e.g. number of eligible patients, number of referrals, etc) and qualitative data collected from interviews with participating providers (e.g. exploring referral process and patient response to care pathways).

2.3 Integrated Care Pathway

The Integrated Care Pathway (ICP) provides both on-site physical therapy (PT) services and centrally-delivered services via telephone or video from study providers at the Durham VA. The ICP is initiated with a primary care referral to the pathway. Patients will start the ICP by receiving on-site PT services for 1-2 visits. These visits will include examination, pain modulation treatment, and patient education. After the on-site visit(s) patients will be referred to receive weekly calls for 6 weeks of home physical activity instruction. These calls will be conducted by providers at the Durham VA. After 6 weeks the patients will be instructed to return to the onsite PT services at participating clinics. Patients then complete a simple 9-question risk stratification tool (SBST) patients scoring 'medium or high risk' on the SBST will receive additional centrally delivered intervention of weekly calls for 6 more weeks from the providers at the Durham VA. This intervention will be tailored to include psychological and behavioral activation components, consistent with what can be delivered by physical therapists. The third step of the ICP will consist of 6 treatment sessions over 6 weeks for those that have been identified as medium or high risk on the SBST.

There will not be strict control of specific interventions delivered in the ICP. However, the ICP incorporates interventions that align with best practice for low back pain management (i.e., non-pharmacological pain treatment, early physical activity, risk stratification, and behaviorally focused intervention). Moreover, inclusion of centrally-delivered services through the Durham VA will help to ameliorate current delays in VA access to services. If there is interest from participating sites, on-site providers will be trained to take over centrally delivered components after the study period ends.

Pathway Procedure for the ICP

The overall flow for the integrated care pathway is described in **Figure 1**, with each step described in more detail in the Figure legend. A TIDier summary[14] for the ICP is also included in **Appendix II**.

focuses our treatment moderation analyses to 1) chronic vs. acute LBP and 2) previous opioid use vs. opioid naïve. This is a embedded pragmatic trial so these two subgroups were selected for their clinical relevance (i.e. can be easily identified during routine encounters) and because their potential as treatment moderators has been established in the literature. Indeed a 2015 meta-analysis on this topic included LBP status and narcotic medication use as potential treatment moderators with strong statistical evidence ($p < 0.05$) supporting their inclusion in confirmatory analyses.[38] We hypothesize that we will observe better outcomes from the ICP for patients with chronic pain and opioid exposure based on previous studies.[38,40,41] Including planned subgroup analyses in this embedded pragmatic trial will complement the primary trial results by providing important information to guide decisions about optimal targeting of resources. Moreover, by focusing on subgroups that have been prioritized in the scientific literature, these analyses also advance the field in this emerging area.

4. SELECTION AND ENROLLMENT OF CLINICAL SITES AND PARTICIPANTS

This study is a 4-year cluster randomized embedded pragmatic trial comparing the effectiveness of two care pathways for increasing non-pharmacological treatment options for LBP. The study target population is veterans seeking services for lower back pain in participating VA clinics. There will be selection of clinical sites and participants for this trial.

Clinical Site Selection

On May 16th we met with VISN6 leadership to distribute an update on the AIM-Back trial and being the process of recruiting interested clinical sites. Formal site agreements will be pursued once the study protocol is finalized (Appendix I).

Unit of randomization: VA primary care clinics.

Sample size: 16 clinics (8 clinics randomized to each care pathway).

Clinic eligibility:

- Volunteer and return signed clinic participation agreement
- Availability of clinical personnel willing to deliver the treatment interventions (in either arm).
- Staff and location need to be distinct from other enrolled clinics.
- Clinics can have a variable number of providers, but together meet criteria for range of monthly visits for LBP

4.1 Participant Inclusion Criteria

At the level of the individual, participant eligibility criteria are intentionally broad in keeping with the pragmatic nature of this trial. Eligibility criteria will be applied equally to and are appropriate for individuals seeking care at any of the clinical sites. Site training will include instruction in the referral criteria so that providers have consistency in the selection of participants for the care pathways (see below). Our intent is to: 1) examine the care pathways in the way they would be used in clinical practice; and 2) maximize diversity and the potential to generalize the findings from this trial.

Referral Criteria (part of site training and used by providers for pathway determination)

1. LBP that is localized or radiating
    - Localized = symptoms reported in the T12 to S5 region only, occurring unilaterally or bilaterally

5. IMPLEMENTATION OF CARE PATHWAYS

VA clinics that have agreed to have their clinical staff trained in offering these care pathways as standard care will be randomly assigned to one of the two care pathways. The manual for training materials used for participating VA clinics is included in Appendix V. VA clinics implementing these back pain care pathways are not engaged in research because both pathways involve different ways of providing standard of care treatment. Employees at clinical sites that implement the back pain care pathways will participate in educational training for delivering the treatment as part of standard care, clinical program-related duties only.

AIM-Back will employ implementation facilitators to work with clinics to implement and maintain their care pathways. AIM-Back facilitators will work directly with a Point of Contact selected by each site to develop procedures for keeping the care pathways and referral procedures top of mind for referring providers, using existing infrastructure at each clinic (e.g. staff meetings). Employees delivering the clinical treatment will not be directly involved in any research related procedures (e.g. follow up survey collection over the phone). To assist each site in implementing the care pathways we will collect data and provide reports to be used for process improvement. These will include both quantitative (e.g. number of eligible patients, number of referrals, etc) and qualitative data collected from interviews with participating providers (e.g. exploring referral process and patient response to care pathways).

5.1 Qualitative Data Collection

In-depth, semi-structured, individual telephone interviews (~20-30 minutes each) with be conducted with a sample of the participating, referring providers. We will purposefully sample approximately 2 referring providers from each clinic as a function of low/high adoption as evidenced by number of referrals at 2.5 to 3 months (based on clinic start dates). Interviews will be conducted as close as possible to the 3-month time frame (+/- 2 weeks to allow for scheduling). A structured form for notetaking will be utilized in order to facilitate rapid qualitative analysis. Information obtained will be used to create case memos (case study qualitative design), which will be fed back to the clinics and triangulated/supplemented with process improvement notes from the facilitator. We will target providing this feedback to the clinics by around the halfway point of the 9 months (i.e., 4.5 months). Information obtained will also inform the next starting clinics with an evaluation across all cases (multiple case study design) to be conducted at the conclusion of the active study/pathway period. The interview guide will be informed by the Consolidated Framework for Implementation Framework (CFIR)[42] and include in-depth, descriptive questions on barriers and facilitators at multiple levels (i.e., provider, clinic, facility, and patient) from the perception of the referring providers. All qualitative data will be used for quality improvement purposes (i.e. will not be considered human subjects research).

5.2 Concomitant Interventions

We have taken steps to minimize the risk of cross-contamination through our choice of study design and by imposing restrictions on clinic participation. That is, clinics must be located at geographically distinct locations, rather than on the same campus. We will measure and report the number of patients with referrals to both LBP care pathways. Furthermore, evidence of concomitant diagnostic testing and treatments received outside of the care pathways will be

We have a solid plan for dropout of a clinic or low recruitment in block 1 that maintains the randomization and balance of covariates with the ability to add clinics to block 2 before the block 2 randomization.

To guard against drop out of clinics from block 2 we plan to add 2 reserve sites to the block 2 wave of randomization to use in case a clinic should drop out of block 2. We have adjusted the block sizes to have 10 clinics randomized in block 1 and have 8 clinics in block 2 (if no clinics drop out in block 1) which include 2 reserve clinics. Through the design process for this trial we have increased the number of clinics from 12 to 16 in part to provide a safeguard for the potential drop out of a clinic and still be able to conduct a successful trial (**see Table 7** below that has effect sizes for 12 clinics with same number of patients per clinic as we targeting for 16 clinics and for 16 clinics with smaller number of patients per clinic).

Table 7. Sample Size Table for AIM-Back

Number of clinics per condition	Number of patients per clinic needed at post-intervention follow-up/ Total N	Df for baseline covariate adjustment	Minimum Detectable Difference (Effect size) ICC=0.01	Minimum Detectable Difference (Effect size) ICC=0.02	Minimum Detectable Difference (Effect size) ICC=0.05	Number needed at baseline with 20% attrition per clinic/ Total N_b
6	42/N=504	5 2	0.49 0.44	0.56 0.50	0.72 0.65	53/ N_b=636
6	84/N=1008	5 2	0.40 0.35	0.47 0.42	0.66 0.59	105/ N_b=1260
8	32/N=512	5 2	0.41 0.40	0.46 0.44	0.57 0.55	40/ Nb=640
8	63/N=1008	5 2	0.32 0.31	0.38 0.37	0.52 0.50	79/ Nb=1264
8	42/N=672	5 2	0.37 0.36	0.42 0.41	0.54 0.53	53/ Nb=848
8	84/N=1344	5 2	0.30 0.29	0.36 0.35	0.50 0.49	105/ Nb=1680

9. STATISTICAL CONSIDERATIONS

9.1 General Design Issues

This study will be a longitudinal cluster randomized trial conducted at 16 VA clinics. The “clinics” must be geographically distinct with no overlap in staffing. A summary of study measures, with their associated definition, data source, and time interval is provided above in Section 10 (Data Collection and Measures). Outcomes derived from VA EHR data (for example utilization and pharmacy data, as well as health factor data obtained through data capture tools designed and implemented for this study) will be obtained for the larger administrative sample (n=1680). Outcomes from telephone administered surveys will be available on approximately half of participants (n=848), who are consented to participate in enhanced survey data collection.

Table 9. Minimum detectable effect size differences by a binary moderator for 90% power, alpha of 0.025 and ICC of 0.01, 0.02, and 0.05 for administrative sample.

Number of clinics per condition	Number of patients per clinic needed at post-intervention follow-up (Total N)	Minimum Detectable Effect Size Difference ICC=0.01	Minimum Detectable Effect Size Difference ICC=0.02	Minimum Detectable Effect Size Difference ICC=0.05	Number of patients per clinic needed at baseline with 20% attrition (Total N_b)
8	21 (N=336)	0.48	0.53	0.63	27 (N_b=432)

9.2.1 Treatment Assignment Procedures

In this study, randomization occurs at the clinic (i.e., the cluster) level and we plan on conducting a covariate constrained randomization[2-4] with the following clinic level covariates assessed prior to randomization: 1) average pain scale scores of LBP patients at clinic, 2) average level of opioid exposure of LBP patients at clinic, 3) number of participating primary care providers at clinic, 4) clinic location (main medical center/community clinic) and 5) average age of LBP patients at clinic. These characteristics were chosen to represent factors likely to be associated with baseline differences in patient population that may affect primary outcomes. However, the exact number of covariates we use will depend on the quality and distribution of these covariates across the 16 clinics that we enroll. For example, if all the clinics we enroll are community clinics, then we would drop the clinic location covariate in the randomization. Currently, throughout the VA, general pain scale scores (rate pain on a scale from 1 to 10) are being collected for all primary care visits and stored in administrative data. Our plan will be to use these data, averaging the pain scale scores of LBP Veterans at each clinic over a 3 to 6-month time period prior to clinic enrollment. For Durham primary care clinics, we found that pain scores were collected at over 95% of primary care visits for lower back pain over a one-year time period. From the Goode et al. pilot study[33] of Veterans with back pain (n=50), we found that correlation of pain scale scores to PROMIS pain interference at baseline was approximately 0.55, a reasonably strong correlation. For the opioid exposure covariate, our plan is to apply the PMC3 opioid exposure definition to LBP patients at clinics over a 3 to 6-month period prior to clinic enrollment and then use the proportion of patients that meet the opioid exposure definition as the average level of opioid exposure for the clinic. The number of participating (referring) providers at a clinic will be used as an indication of size of clinic (volume of patients served). For the age of LBP patients, we will use age that is available in the electronic health record and average the age of Veterans with LBP visits to the clinic over the 3 to 6-month time period prior to clinic enrollment.

In planning for the UH3, we have determined that we cannot roll out the interventions for all 16 clinics at the same time due to logistical constraints. Therefore, we plan to conduct a covariate constrained randomization in 2 blocks following the extension to randomize multiple blocks of groups by Carter and Hood.[54] In the first block we will have 10 clinics and the second block will have 8 clinics if no clinics drop out of block 1 before randomization occurs in the second block. The second block will include 2 reserve clinics for use if a clinic should drop out from block 2 (see Section 8). In the standard covariate constrained randomization all participating clusters need to be enrolled before randomization. We are concerned that clinics that would roll out later in the schedule may lose interest or drop out if we enroll and randomize all clinics at once at the beginning of the study. An advantage of the blocked compared to the standard randomization is that, if a clinic would drop out in the first block before we randomize the second block, we would

β_1*(moderator) + β_2*(moderator*intervention)+ β_3*(followup) + β_4*(followup*intervention) +β_5*(moderator*followup) +β_6*(moderator*intervention*followup) for clinic *i*, patient *j,* at time *k*. A random effects (clinics and time by clinics) will be included in the model to account for clustering of patients within clinics, as the clinics are the unit of randomization. Differential change in pain interference or function by moderator for the two intervention groups is supported if the coefficient for β_6 is significantly different than zero at the p=0.025 level. For the adherence multinomial outcome, we will fit a generalized linear model with a random effect for clinic and cumulative logit link function. As the adherence outcome is potentially a 3-level outcome there will be two interaction terms to examine for moderation effects (e.g. whether the effect of the intervention by chronicity is different for those moderately adherent vs. not adherent and/or those fully adherent vs. not adherent).

Exploratory analyses to identify multidimensional subgroups from combinations of baseline characteristics (e.g. age, sex, opioid use) for treatment moderation of the primary outcomes of pain and function as well as per protocol adherence to treatment.

Patients are likely to vary in treatment response. This variation is known as heterogeneity of treatment effects (HTE). Exploration of HTE has typically focused on one-factor-at-a-time post-hoc subgroup analyses (as described in our planned subgroup analysis); however, it is possible to instead identify multidimensional subgroups exhibiting HTE. In particular, we will apply the data-driven method GUIDE[66](Generalized, Unbiased, Interaction Detection and Estimation) that can be used for longitudinal outcomes, so can incorporate all follow-up outcomes at 3 and 12-months, as well as multi-response outcomes. GUIDE uses a regression tree algorithm that identifies subgroups with heterogeneous effects and estimates how the treatment effect varies across the subgroups. We will include baseline demographic (e.g. age, sex, race) and clinical characteristics (e.g. LBP chronicity, opioid use, comorbid conditions) that are available from EHR. GUIDE provides bootstrap confidence intervals for the treatment effects of identified nodes and is available via a software package.

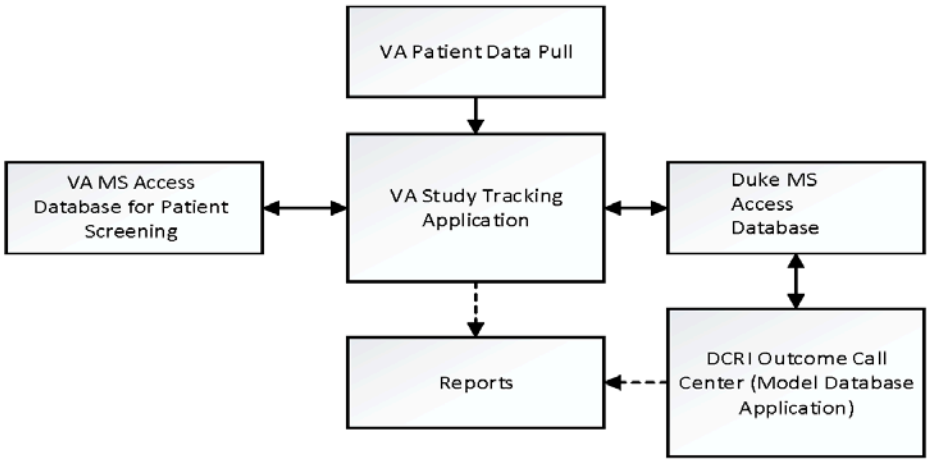
We will also examine “adherence” among all patients referred to one of the clinical pathways. We define “adherence” according to participant participation in planned sessions with intervention team personnel (telephone or in-person) in each care pathway. The number of planned sessions will vary between and within arms (e.g. depending on response to 1st service in pain navigator arm and depending on risk stratification in integrated pathway arm); therefore, we will define adherence as follows: 1) non-adherent – attended no planned sessions 2) partially adherent – attended some, but not all planned sessions 3) attended all planned sessions. Depending on the number of planned sessions in each pathway we will explore additional multi-level variables that may be more informative; structure of the adherence variable for Aim 2 analyses will depend on data distribution.

We will describe adherence in both care pathways in order to identify patient characteristics potentially associated with program participation and improved pain outcomes. We will not test for differences in adherence rates between pathways as our goal is to gain insight into where to target clinical programs and focus efforts to improve access to and patient engagement with non-pharmacologic pain services. We will also conduct exploratory analyses on alternative definitions of adherence or aspects of intervention receipt. For example, in the pain navigator arm, differentiating based on whether someone actually received one of the recommended services.

Exploratory analyses to identify qualitative interactions that define groups of Veterans that have greater improvement in pain and function with the integrated care pathway versus the care

in a secure Microsoft Access database at Duke. The recruitment and participant management system will be a Durham VA in-house application named Study Tracking that will be run by the VA staff. It will provide all functionality needed to keep track of study participants as they move through the protocol. Study Tracking will interface with the Microsoft Access database. Following consent, participant contact information will be duplicated from Study Tracking into the designated MS Access database at Duke. DCRI's Call Center software will interact with the MS Access database to create a call schedule. When someone is due for a follow-up call, a member of the Call Center will be prompted to contact the participant and the follow-up survey results will be stored in MS Access. **Figure 5** provides a general overview of the data flow for the project across the two participating entities.

Figure 5- General Data Flow Overview for AIM-Back



10.2 Data Management

The Durham VA and Duke Clinical Research Outcomes Operations group will serve as the coordinating center and data management center for this study. The Durham VA will be responsible for enrollment, consenting participants for the follow up telephone survey assessments, baseline data collection, as well as data analysis after the database lock. The DCRI call center will be responsible for administering follow up surveys. Other than standard demographic information, all of the outcome measures being collected for this study via follow up surveys are listed in the appendix and are harmonious with other existing patient reported pain and pain impact measures collected within the field.

The follow up surveys will be administered at three, six, nine, and twelve months post baseline, and participants will be paid $25 for completing each survey, including baseline. Payments will be made via online gift card or mailed to the participants' address. Payments will be administered through the Duke Clinical Research Institute Call Center.