MIDAS Quality Enhancement Research Initiative
Center for Clinical Management Research
2215 Fuller Rd (152)
Ann Arbor, MI 48109
(734) 845-3502

Official Title of Project: MIDAS Cluster Randomized Controlled Trial of Implementation Strategies to Optimize Use of Medications in VA Clinical Settings (MIDAS cRCT)

ClinicalTrials.gov ID: NCT05065502

Date: 09/24/2025

Please find the Study Protocol and Statistical Analysis Plan below. Since the creation of this document, the following changes have been made to the analysis plan:

- Data was collected through 18 months post-baseline rather than through 36 months due to time constraints.
- Instead of a patient-level dichotomous response over time, the differences between arms were modelled as a single outcome adjusting for the evidence-based practice (EBP) and the level of potentially inappropriate medications (PIMs) in the pre-period to address the large difference in PIMs between EBPs.
- After conducting additional testing of our algorithm for detecting CBTI sessions, we found the algorithm to be adequately sensitive to detecting *any* CBTI but not accurate for detecting the number of CBTI sessions (e.g., CBTI could often be mentioned in the context of a referral or some other ongoing treatment). Thus, the secondary outcome “Change in Mean Cognitive Behavioral Therapy for Insomnia (CBTI) Sessions Completed” could not be analyzed.
- We had originally planned for all CBTI sites to provide information regarding where in the medical records we could identify referrals. However, prior to launching the trial, we determined several of the sites' processes for referrals were not distinguishable as referrals vs. other forms of communication (e.g., adding a therapist as a cosigner to a note), making collection of this data impractical. Thus, we did not have sites collect referral criteria and the secondary outcome “Change in the Monthly Percentage of Patients Referred to Cognitive Behavioral Therapy for Insomnia (CBTI) Across Facilities” could not be analyzed.

A protocol paper for this project, titled “Maintaining Implementation through Dynamic Adaptations (MIDAS): protocol for a cluster-randomized trial of implementation strategies to optimize and sustain use of evidence-based practices in Veteran Health Administration (VHA) patients,” was published in May 2022.

Jeremy Sussman, MD, MS
Principal Investigator

Aims

Mixed methods analyses will be used to evaluate the following three aims:

1. *To compare the effectiveness of two implementation strategies (LEAP Quality Improvement Learning Program + Academic Detailing (AD) vs. AD alone)* on potentially inappropriate medication use, using a pooled analysis of effects across three trials at 18 months, two years, and three years post-baseline at the clinic-level, based on monthly assessed data from 13-36 months*;*
2. *To compare the effectiveness of the two implementation strategies* on secondary outcomes specific to each trial at 18 months, two years, and three years post-baseline, based on monthly assessed data from 13 to 36 months; and
3. *To explore the effects* of implementation, provider behaviors and experiences, and context, on sustained improvements in potentially inappropriate medication use.

Methods

Clinic Randomization

Within each trial, clinics will be randomized after assenting to participate (equivalent to enrollment). Clinics will be assigned to one of two arms by a statistician, stratified further by clinic type (medical center, community clinic, or Community Living Center) if needed to ensure partial balance between arms with respect to potential confounders associated with culture and complexities associated with clinic location [1,2].

Outcomes and Analyses

As part of a pooled analysis, we will compare the same two implementation strategies across all three EBPs and take a unified approach to implementation and evaluation across the trials. Table 1 shows MIDAS measures, data collection timeframe, and data sources. While a unified dichotomous outcome, i.e., potentially inappropriate medications, was identified for each trial to allow for the pooled analysis, each trial will also be analyzed individually (see Table 1).

Table 1: MIDAS Measures showing Data Sources and Timepoints by Aim

Aim	Type	Source
Aim 1: Primary outcome		
Service Outcomes		
Proportion of potentially inappropriate care	Quantitative	CDW administrative data
Aim 2: Secondary outcomes		
Service Outcomes		
VIONE Trial		
Potentially inappropriate use of proton pump inhibitors (PPIs), aspirin, and CNS active medications (e.g., muscle relaxants), analyzed one at a time	Quantitative	CDW administrative data
Monthly medication costs for all drugs	Quantitative	CDW administrative data
Number of pharmacist medication reviews	Quantitative	CDW administrative data

Aim	Type	Source
DOAC Trial		
Dashboard flags including potential mis-dosing, potential medication interactions, or concern for nonadherence	Quantitative	CDW administrative data
CBTI Trial		
Prevalence of any CBTI receipt	Quantitative	CBTI note templates completed by CBTI therapists
Patient Outcomes		
VIONE Trial		
Number of inappropriate medications by patient		CDW administrative data
Aim 3: Exploratory outcomes		
Employee Behavioral Outcomes		
Continuous Quality Improvement Assessment	Quantitative	LEAP and AD participants
Workgroup Cohesion & Engagement Scale	Quantitative	LEAP and AD participants
QI skill application	Quantitative	LEAP participants
Employee Experience		
Burnout	Quantitative	LEAP and AD participants
Best Places to Work	Quantitative	LEAP and AD participants
Process Evaluation		
Pragmatic Context Assessment Tool	Qualitative/ Quantitative	LEAP participants
Semi-structured interviews	Qualitative	Purposive sample of LEAP and AD participants
How often the provider uses the dashboard	Quantitative	CDW administrative data
Rates of new DOAC starts compared to warfarin starts	Quantitative	CDW administrative data
Fidelity	Qualitative/ Quantitative	Academic Detailer and Champion/LEAP Coaches
Provider satisfaction	Qualitative/ Quantitative	LEAP and AD participants
Intentions	Qualitative/ Quantitative	LEAP participants
Detailing visit documentation	Qualitative/ Quantitative	Academic Detailer and Champion
Coaching documentation	Qualitative/ Quantitative	LEAP Coaches
Semi-structured interviews	Qualitative	Purposive sample of LEAP and AD participants

Aim 1: Primary Outcomes and Pooled Analysis.

Although each trial will be conducted as an independent study, our primary aim is to compare across trials the effectiveness between the two implementation strategy arms in reducing potentially inappropriate medications during post-implementation period. To this end, we defined a unified primary outcome to allow us to combine the results *across* the three trials. The unified primary outcome will be operationalized based on a patient-level dichotomous response indicating potentially inappropriate

medication use (yes/no) among patients at-risk of potentially inappropriate medications, i.e., among those who may benefit from the specific EBP each month. The monthly patient-level potentially inappropriate medication use response will be summarized to clinic-level month-by-month percentage of potentially inappropriate use using administrative data from baseline to 36 months, with months 13-36 as the post-implementation follow-up period. Each trial-specific monthly data will be cross-sectional, i.e., different patients may be included in each month.

For *inappropriate polypharmacy*, the clinic-month outcome will be the proportion of patients who had medication possession (based on VA pharmacy fill data) of one or more medications from the AGS Beers criteria that are included on the VIONE PIMs dashboard [3] (numerator) among patients age 65 or older, not receiving palliative care, and followed by the clinic (denominator). For each drug included on the PIMS dashboard, there are associated business rules that define when medication use is flagged as potentially inappropriate; these same criteria will be applied in this trial. For example, use of a first or second generation anti-psychotic drug is flagged as potentially inappropriate unless there is a diagnosis of schizophrenia or bipolar disorder. These criteria had previously been determined by VIONE's Subject Matter Expert group, which provides VIONE with guidance on translating deprescribing criteria into the most practical and appropriate rules for use on the dashboard. Altogether, the following AGS Beers medications from the PIMS dashboard will be included in the analysis: anticholinergics, antipsychotics, aspirin, benzodiazepines, long-acting sulfonylureas, muscle relaxants, NSAIDs, proton pump inhibitors (PPIs), sliding scale insulin, and Z-drugs.

For *DOAC safety*, the outcome will be the proportion of patients with potentially inappropriate prescribing out of those using DOACs, as measured by "flags" (e.g., potential mis-dosing based on renal function and other indicators) on the DOAC dashboard. The DOAC flagging system is based on FDA indications and has been in clinical use since 2018. Components of the outcome include inappropriate dosing for the given indication and use of DOACs in contraindicated settings (such as valve replacements).

For *first-line treatment for insomnia*, the outcome will be the proportion of patients with a new prescription for a sedative-hypnotic medication who have not had CBTI in the prior 12 months out of all primary care patients actively following with the clinic and are not in hospice/palliative care.

For all three trials, medication use (yes/no) and possession of active prescription for each month will be determined using exposure days based on supply days, and use will be determined by the exposure status on day 1 of each month. We will also do sensitivity analyses based on the criteria of use anytime during the month as well as PIMs defined to medications used chronically, for example, greater than 90 of the 180 prior days.

For each trial we will first compare demographic characteristics (age, sex, and race) of patients at risk of potentially inappropriate medications in the first month of implementation between the two arms. We will then obtain, for each trial by arm, crude monthly percentages (along with the corresponding 95% confidence intervals) of potentially inappropriate medications, averaged across clinics randomized to each arm and weighted by clinic-month size. For each trial, we will plot the monthly clinic level percentages over the follow-up 13-36 months to graphically assess if the difference between the two arms can be meaningfully summarized across the three trials with the unified outcome. If we find, for example, that trends between-arms over post-implementation months differ notably across the three trials, unified results comparing AD+LEAP vs. AD arm across trials may not be meaningful, and we will only conduct analyses separately by each trial.

For comparison between arms, we will use generalized estimating equations (GEE) with clinic level monthly percent of potentially inappropriate medications among patients at risk during post implementation period (months 13 to 36) as the dependent variable. The model will include indicators of two trials with one trial as the referent category to account for differing underlying levels of inappropriate medication use across trials. The model will also include follow-up time in months and the LEAP+AD arm indicator with AD as the referent category and will adjust for serial correlation within clinic over time. We will also include time by arm interaction to assess if the magnitude of the difference between LEAP+AD vs. AD changes over time. If the interaction is significant, we will estimate between-arm difference at 18 months as well as at 2- and 3-years separately based on the model with the interaction term. On the other hand, if the interaction is not significant, this would indicate between-arm difference not to differ at the three follow-up times of interest (18, 24, and 36 months), and thus we will drop the interaction term and the parameter estimate of the LEAP+AD arm indicator will be used to estimate the time-averaged difference in percentage of patients with inappropriate medications during post-implementation period in clinics randomized to LEAP+AD compared to clinics randomized to AD.

If we find notable baseline demographic difference between arms within trials, we will use a generalized linear mixed model (GLMM) with logit link to estimate the between-arm difference while adjusting for baseline age, sex, and race difference with monthly person level response (yes/no) data from post-implementation period of months 13 to 36. In addition to time, AD+LEAP indicator and trial type indicators as predictors, the GLMM model using patient level data will include patient age, sex, race, and random intercepts for patients nested within clinic to adjust for potential correlation within clinics and serial correlation over time. The parameter estimate for the LEAP+AD arm indicator will be used to estimate the time-averaged odds of inappropriate medication use during the post-implementation period for patients in clinics randomized to LEAP+AD compared to the odds of the same patients if their clinics were randomized to AD. Although the GEE and GLMM models give different summary estimates with different interpretations, the GLMM model allows for adjusting for patient characteristics, and a consistent substantive conclusion will assure us of the evidence for the effect of LEAP when added to AD. Similar to the GEE model, we will test if the odds ratio of LEAP+AD vs. AD changes over time by including time by arm interaction term, and if the interaction term is significant, we will obtain adjusted ORs associated with LEAP+AD compared to AD at 18 months, 2-years, and 3-years.

For each trial we will also compare AD and AD+LEAP to usual care controls. To do this, we will perform a non-randomized secondary analysis for each trial. The analysis will have the same primary outcome variable and use the same generalized linear mixed model (GLMM) with logit link. The primary control group will be all non-participating sites. We will also use a secondary analysis, where for each intervention site we will have two control sites that are matched on clinic size (within 50%), pre-intervention outcome rate (within 30 rankings of all sites), and region of the country. These analyses will adjust for the clinic-level variables clinic size, intervention outcome rate, region of the country and the patient-level variables age, sex, and race.

Aim 2: Secondary Outcomes and Analyses

Secondary outcomes for VIONE will be the prevalence of potentially inappropriate use of PPIs; the prevalence of potentially inappropriate use of aspirin; and the prevalence of potentially inappropriate use of CNS active medications (muscle relaxants, anti-psychotics, Z-drugs, and

benzodiazepines) or anticholinergic drugs; number of inappropriate medications at a patient level; monthly medication costs for all drugs, without regard to appropriateness; and number of pharmacist medication reviews.

Secondary outcomes for the DOAC trial will be the sub-components of the "flags" on the dashboard. These include potential mis-dosing, potential medication interactions, or concern for nonadherence. This follows the organizational structure of both the presentation of the flags on the dashboard and the key messages provided to the academic detailing and LEAP teams. Process outcomes will be how often the provider uses the dashboard and rates of new DOAC starts compared to warfarin starts. These outcomes will be kept in alignment with our other work using the dashboard [4].

In stand-alone analyses of the CBTI trial, the primary outcome will be the prevalence of any CBTI receipt among primary care patients actively following with the clinic who are not in hospice/palliative care.

Analyses of secondary outcomes such as percent of potentially inappropriate use of PPI or mean number of CBTI sessions at each clinic month will be similar to that of the primary outcome using the GEE model accounting for correlation over time. We will also conduct separate analyses by trial with the dependent variables that are unique to each trial. For example, for the polypharmacy trial, secondary outcome of interest is count of medications flagged as inappropriate based on Beers' criteria [3]. We will compare monthly rates of Beers' list medication use between implementation strategies using generalized linear mixed models with log link.

Aim 3: Exploration of Potential Predictors of Clinical Outcomes

Our process evaluation will follow a multi-phase concurrent nested mixed methods design [5]. This design has three purposes: 1) help prepare all stakeholders and participants prior to the start of each trial; 2) monitor progress of implementation; and 3) explain summative findings. Overall priority is placed on quantitative methods that guide the trials, while qualitative methods are embedded or "nested" within conduct of the trials.

Employee behavior and experience measures will be collected via five scales as listed in Table 1. Surveys will be administered via online link within invitation emails; administration will occur at baseline and 12-months post-baseline; satisfaction will be elicited at the end of each intervention (upon completion of the 6-month "core" LEAP program for LEAP team members and at the end of each AD visit for AD participants). Descriptive statistics will be generated and tests for differences across implementation strategy arms will be conducted using mixed models to account for within-clinic correlation.

Qualitative data will be collected prior to and 18-months following baseline via semi-structured interviews (virtual by telephone or conferencing software (e.g., MS Teams platform)). A purposive sample of key people (clinic leaders, supervisors, providers, and staff) at each clinic will be invited to participate so we can better understand the context in which the implementation strategies are/were deployed. The interview guides and qualitative analyses will be guided by the CFIR to identify potential and actual barriers and facilitators [6–8]. Principles embedded within the DSF will guide exploration of the degree of engagement in quality improvement and teamwork [9]. Prior to implementation, this information will help inform the work of the academic detailers and LEAP Coaches; post-

implementation, this information will help to explain quantitative findings within and across the trials. Interviews will be audio recorded and transcribed verbatim. Pre-implementation, interviews will focus on collecting practical information using a rapid analysis approach [10,11] to help tailor and adapt implementation for each participating clinic. Post-implementation, qualitative analyses will seek insights on what kinds of improvements were made, barriers and facilitators to making improvements, reflections on/satisfaction with participation in AD/LEAP and explore *relationships* between determinants, participants, and key stakeholders and how these may lead to building coalitions of support [12,13]. We will combine qualitative findings with quantitative measures from Aims 1 and 3 to help explain changes (or lack of) over time.

Our process evaluation will rely on quantitative and qualitative data sources. Fidelity to each implementation strategy will be tracked by interventionists (the detailers and LEAP coaches) completing a mixed-methods self-assessment tool after each interaction (a coaching session for LEAP, detailing contact for AD). These assessments will be used to guide coach-supervisor and peer reflections on improvements, problem-solving, and mitigating barriers and amplifying facilitators of improvement efforts. We will also track participation by participants (individuals scheduled for detailing and/or LEAP team members) and completed assignments by LEAP teams. The Academic Detailer and LEAP coaches will enter notes for each interaction into a tracking system for each strategy. This data will be combined with pre-implementation and 18-month semi-structured interview data for further insights into barriers, facilitators, and problem-solving approaches used by LEAP coaches and detailers. Quantitative and qualitative data will be combined at the analysis or interpretation phase.

Economic Evaluation

We will use a micro-costing method [14,15] to determine the costs to deliver LEAP and academic detailing (AD). The LEAP coaches and academic detailers developed a list of the activities they will perform for each participating site. Depending on the specific activity, they determined the best way to record the time spent on each activity—e.g., logging the start and stop time each time the activity takes place vs. setting an estimated average time for activities that take approximately the same amount of time for each incidence (such as recurring meetings, responding to quick queries via e-mail, meeting preparation, etc.). In the latter case, the coaches and detailers simply record the occurrence of the activity, which is then assigned the estimated time. The coaches and detailers will log times for each activity, categorized by participating site, in a time tracking database. Using data from this database, we will calculate the average time required for each activity and apply this to the number of times it takes place over the course of performing the implementation strategy at a site. These data can then be used to determine an estimated total time required to perform the implementation strategy (LEAP or AD) at a site, which, combined with the hourly cost of the LEAP coach or detailer, can be used to calculate the total cost of employing the strategy at a site.

List of Abbreviations

Academic detailing (AD)

American Geriatrics Society (AGS)

Cognitive Behavioral Therapy for Insomnia (CBTI)

Department of Defense (DoD)

Direct-acting anticoagulant medications (DOACs)

Dynamic Sustainability Framework (DSF)

Evidence-based practices (EBPs)

Generalized estimating equations (GEE)

Generalized linear mixed model (GLMM)

Learn. Engage. Act. Process. (LEAP)

Maintaining Implementation through Dynamic Adaptations (MIDAS)

Massive-Open Online Course (MOOC)

Office of Mental Health Services and Suicide Prevention (OMHSP)

Plan-Do-Study-Act (PDSA)

Quality Enhancement Research Initiative (QUERI)

Quality improvement (QI)

VA Center for Medication Safety (VA MedSAFE)

Veteran Health Administration (VHA)

Veterans Affairs (VA)

References

[1] Hayes RJ, Moulton LH. Cluster randomised trials. 2009.

[2] Campbell MJ, Walters SJ. How to design, analyse and report cluster randomised trials in medicine and health related research. 2014.

[3] By the 2019 American Geriatrics Society Beers Criteria® Update Expert Panel. American Geriatrics Society 2019 Updated AGS Beers Criteria® for Potentially Inappropriate Medication Use in Older Adults: 2019 AGS BEERS CRITERIA® UPDATE EXPERT PANEL. J Am Geriatr Soc 2019;67:674–94. https://doi.org/10.1111/jgs.15767.

[4] Barnes GD, Sippola E, Dorsch M, Errickson J, Lanham M, Allen A, et al. Applying population health approaches to improve safe anticoagulant use in the outpatient setting: the DOAC Dashboard multi-cohort implementation evaluation study protocol. Implementation Sci 2020;15:83. https://doi.org/10.1186/s13012-020-01044-5.

[5] Creswell JW, Plano Clark VL, Gutmann ML, Hanson WE. An Expanded Typology for Classifying Mixed Methods Research Into Designs. The Mixed Methods Reader, Sage Publishing; n.d.

[6] Damschroder LJ, Lowery JC. Evaluation of a large-scale weight management program using the consolidated framework for implementation research (CFIR). Implementation Science 2013;8:51.

[7] 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;7:233–41. https://doi.org/10.1007/s13142-016-0424-6.

[8] Damschroder LJ, Reardon CM, AuYoung M, Moin T, Datta SK, Sparks JB, et al. Implementation findings from a hybrid III implementation-effectiveness trial of the Diabetes Prevention Program (DPP) in the Veterans Health Administration (VHA). Implementation Science 2017;12:94.

[9] Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. Implement Sci 2013;8:117. https://doi.org/10.1186/1748-5908-8-117.

[10] Averill JB. Matrix Analysis as a Complementary Analytic Strategy in Qualitative Inquiry. Qual Health Res 2002;12:855–66. https://doi.org/10.1177/104973230201200611.

[11] McMullen CK, Ash JS, Sittig DF, Bunce A, Guappone K, Dykstra R, et al. Rapid Assessment of Clinical Information Systems in the Healthcare Setting: An Efficient Method for Time-pressed Evaluation. Methods Inf Med 2011;50:299–307. https://doi.org/10.3414/ME10-01-0042.

[12] Lennox L, Maher L, Reed J. Navigating the sustainability landscape: a systematic review of sustainability approaches in healthcare. Implementation Sci 2018;13:27. https://doi.org/10.1186/s13012-017-0707-4.

[13] Lennox L, Doyle C, Reed JE, Bell D. What makes a sustainability tool valuable, practical and useful in real-world healthcare practice? A mixed-methods study on the development of the Long Term Success Tool in Northwest London. BMJ Open 2017;7:e014417. https://doi.org/10.1136/bmjopen-2016-014417.

[14] Gehlert E, Jacobs J, Barnett P. Costing Methods Used in VA Research, 1980-2012: Technical Report 32 2016.

[15] Gold HT, McDermott C, Hoomans T, Wagner TH. Cost data in implementation science: categories and approaches to costing. Implementation Sci 2022;17:11. https://doi.org/10.1186/s13012-021-01172-6.