Participatory System Dynamics vs Audit and Feedback: A Cluster Randomized Trial of Mechanisms of Implementation Change to Expand Reach of Evidence-based Addiction and Mental Health Care

NCT number: NCT04356274

Version date: 11/05/2021

SPECIFIC AIMS

Providers, patients, policy makers and scientists each have a stake in ensuring all patients with opioid use disorder (OUD), alcohol use disorder (AUD), depression and PTSD, receive timely, evidence-based care. In Veterans Health Administration (VA) these common, costly conditions comprise the majority of outpatient addiction/mental health needs.[1–3] But, highly effective treatments, such as evidence-based psychotherapy (EBPsy) and evidence-based pharmacotherapy (EBPharm) reach only 3-28% of patients.[4–6] National evidence-based practice (EBP) dissemination programs,[7–11] policies,[12–14] incentivized quality measures,[15,16] and EBP-focused electronic health records[17] have been insufficient for greater EBP reach,[18–20] key to preventing chronic impairment, relapse, overdose,[21–23] and suicide[24–26] among Veterans.

In our R21, we piloted participatory system dynamics (PSD) due to the ineffectiveness of status quo Audit and Feedback (AF) for achieving desired EBP reach. Like most health systems,[27–29] VA enlists a multicomponent strategy of team huddles and AF to coordinate and improve EBPsy/EBPharm care. But, AF outcomes vary widely,[30] with the least success for complex tasks, such as multidisciplinary EBP continuity and coordination.[31] In fact, AF effect sizes in research studies have not increased for 14 years.[32] AF and PSD are rich comparators, due to similarities (data review) and differences (system simulations).[33,34] PSD research shows that fixing complex, multidisciplinary, team EBP delivery requires addressing system causes.[35–40]

PSD theory of change is that participatory learning from modeling[41,42] enables new capacities for managing causal system dynamics or mechanisms of change.[43,44] PSD learning to develop systems thinking overcomes cognitive limitations[45,46] to prospectively restructure delay and feedback dynamics that cause EBP reach over time.[47] We iterated with stakeholders to co-create 4 universal models of 'limited EBP reach' (Care Coordination, CC; Medication Management, MM/EBPharm; Psychotherapy, Psy/EBPsy; Aggregate, Agg), understood by everyone to explain reach in *any* team as a function of local resources.[48,49] We extracted existing health system data using standard definitions for diagnoses, appointments, visits, and EBPs to calculate model parameter values for *local teams*.[15,50] Models define mechanisms identified by stakeholders' using systems theory and calculus (structural validity), validated against historical data (structural-behavioral validity).[51] PSD learning is via real-time visual emergence of local causes of EBP reach in virtual experiments.[52–54] Simulations are a safe way to build systems thinking in to EBP decisions.[44,55,56]

Using PSD and AF with outpatient mental health teams, we will achieve the following specific aims:

We propose a two-arm, 24-site (12 clinics/arm) cluster randomized trial to test for superior EBP reach for depression, PTSD, OUD, and AUD (4 EBPsy; 7 EBPharm, see *Plan*) in PSD clinics over AF (aim 1). We test whether PSD works via systems thinking as predicted by PSD theory of change (aim 2). We test whether EBP causal mechanisms in MM/EBPharm and Psy/EBPsy models generalize in AF and PSD clinics (aim 3). Eligible clinics will be from regional VA health care systems (analyses control for clinics nested within VA[57]) below the overall VA national quality median for EBPsy/EBPharm. Computer-assisted stratified block randomization will balance arms (6 clinics/wave over four waves) at baseline using VA data.

Aim 1: Test superiority of PSD over AF for increasing EBP initiation and course. The proportion of patients (1a) initiating EBPsy/EBPharm, and (1b) completing an adequate EBPsy/EBPharm course, will significantly increase in PSD clinics as compared to AF clinics, in omnibus tests across EBPs for (1a) initiation, and (1b) course, using clinic 12 month pre/post period EBP reach averages.

Aim 2: Test PSD theory of change that increased EBP reach is via systems thinking.[47,58,59] The effect of PSD/AF on 12 month period EBP reach will be explained by 6 month team systems thinking (STS), in ratio of mediator probability weight (RMPW) mediation analyses using the R 'MultisiteMediation' package.[60–62]

Aim 3: Test the generality of mechanisms of change in EBPsy/EBPharm PSD models. Structural-behavioral validation tests[51,63,64] of causal dynamics formulated in our Psy/EBPsy and MM/EBPharm models will generalize to explain EBP reach as a function of local data across PSD and AF clinics (in *Plan*).

Exploratory aim. We contextualize aims 1-3 using *provider surveys* of the clinic-level learning organization survey (LOS-27),[65] team-level decision-making (TDMQ)[66] and burnout (PACT),[67–70] and PSD/AF feasibility, acceptability and appropriateness (FIM, AIM, IAM).[71] We enlist *qualitative/observational coding* of PSD/AF fidelity, systems thinking[72] during team decision-making, and online PSD/AF use and sustainment.[34,73]

Innovation – Our design tests a well-established systems science strategy (PSD) for its superiority, theory of change, and causal mechanisms, against a highly prevalent, but variably effective implementation strategy (AF) in the largest integrated health care system in the US.[74] Achievement of specific aims will advance scientific knowledge regarding the causes of limited EBPsy/EBPharm reach. PSD innovates in empowering frontline teams of providers to understand and prospectively manage local causal dynamics to better meet the high-priority needs of addiction and mental health patient populations.

RESEARCH PLAN

SIGNIFICANCE. Current strategies are insufficient for improving reach of evidence-based practices (EBPs). Ten years of VA nationwide programs to disseminate, train, track and incentivize evidence-based psychotherapy (EBPsy) for PTSD and depression patients, and evidence-based pharmacotherapy (EBPharm) for patients with depression, opioid use disorders (OUDs) and alcohol use disorders (AUDs) has been insufficient for reaching adequate proportions of these patient populations (see *Table 1*). Opiate and alcohol misuse, depression and PTSD, are the primary reasons Veterans seek outpatient VA addiction and mental health care.[1,3] Yet, at the median (see *Table 1*), less than 60% of PTSD patients, and less than 40% of depression patients, start psychotherapy of *any kind.* Only 28% of depression patients starting EBPharm receive a therapeutic course, and 71% of OUD patients do not initiate EBPharm. Among depression and PTSD patients who start psychotherapy, only 30-44% are retained for at least 3 visits. *Only 3-5%* of depression and PTSD patients start EBPsy (see *Table 1*).

Table 1. VA SAIL Quality Measures & R01 Study EBP Reach (Q1 2017)

Population Coverage - Denominator (*diagnostic* cohorts)	**Study**	**EBP**	**SAIL NAME**	**VA 50th %ile**
% OUD ***diagnosed*** patients receiving opioid agonist or antagonist	**A1, 2, 3**/recruit	**EBPharm**	SUD16	28.4
% AUD ***diagnosed*** patients receiving medication assisted therapy*	**A1, 2, 3**/recruit	**EBPharm**	ALC-TOP	9.8
% Depression ***diagnosed*** patients with depression psychotherapy visit	recruit		Psy32	38.3
% PTSD ***diagnosed*** patients with psychotherapy visit for PTSD	recruit		Psy38	55.8
Continuity of Care - Denominator (diagnosis + active *treatment*)				
% Patients on ***new antidepressant medication*** (84 days continuous)	**A1, 2, 3**/recruit	**EBPharm**	MDD43h	73.0
% Patients on ***new antidepressant medication*** (180 days continuous)	**A1, 2, 3**/recruit	**EBPharm**	MDD47h	57.2
% Depression ***treatment*** patients - 3 psychotherapy visits in 6 weeks	recruit		Psy33	30.4
% PTSD ***treatment*** patients - 3 psychotherapy visits in 6 weeks	recruit		Psy39	44.0
EBPsy Reach Targets – Denominator (diagnosis/EBP template)				**Reach %**
PTSD - EBP Template for PE or CPT Session 1 (Initiate)	**A1, 2, 3**	**EBPsy**	PTSD 56	5.3

PTSD - EBP Template for PE or CPT Completion (Course)	**A1, 2, 3**	**EBPsy**	PTSD 56	-
Depression - EBP Template for CBT-D, ACT-D, IPT-D Session 1 (Initiate)	**A1, 2, 3**	**EBPsy**	n/a	2.6
Depression - EBP Template for CBT-D, ACT-D, IPT-D Completion (Course)	**A1, 2, 3**	**EBPsy**	n/a	-

Source. SAIL, Corporate Data Warehouse (CDW). EBP = Evidence Based Practice Reach Outcome. A1 = Aim 1. EBPharm = Evidence-based Pharmacotherapy. EBPsy = Evidence-based Psychotherapy. *Psychotropic Drug Safety Initiative (PDSI) Medication Assisted Therapy for AUD = Acamprosate, Disulfiram, Naltrexone, Topiramate. PE = Prolonged Exposure. CPT = Cognitive Processing Therapy. CBT = Cognitive Behavioral Therapy. ACT = Acceptance and Commitment Therapy. IPT = Interpersonal Process Therapy.

Background of the 'Limited EBP Reach' Problem: Ineffective Audit and Feedback (AF) Strategies. VA/DOD clinical practice guidelines, national EBP training and consultation programs, EBP note templates in electronic health records, and VA quality measures known as Strategic Analytics for Improvement and Learning (SAIL),[12,13,15,75,17] were all developed, due to the efficacy of EBPsy and EBPharm (see *Table 1*) for reducing PTSD[76,77] and depression,[78–82] alcohol[83–86] and opioid misuse,[87,88] and reducing risk of death by suicide or overdose.[23,26] VA dissemination programs also demonstrate EBPsy/EBPharm effectiveness. Patients who received cognitive behavioral therapy (CPT) for depression experienced a 40% reduction in depression symptoms[8] and over 60% of Veterans who received prolonged exposure (PE) experienced clinically significant improvements in PTSD.[10] VA's **multicomponent strategy** of 1) setting guidelines, 2) providing training/resources to meet guidelines, 3) tracking performance, and 4) providing feedback to identify gaps or progress, form the chief principles of AF benchmarking in any health system.[15,16,29] But, AF systemic reviews show AF must also provide actionable insights to gain increases in EBP reach.[27,31,32,89–91]

Reach. We define **reach** as the proportion of the outpatient population who *initiate* an EBP and complete a therapeutic *course*. Implementation gaps for depression, PTSD, AUD and OUD patients (*Table 1*), reflect the complexity of identifying optimal health system improvements. Too few VA patients receive EBPsy even in specialty programs.[11,92] Improving reach involves *interdependent fixes* within and across programs (e.g., general vs. specialty), meeting a variety of common, and often comorbid patient needs (e.g., mental health/addiction), by multidisciplinary teams with varying capacities/expertise (e.g., EBPharm prescribers vs. EBPsy providers). ***What generally causes limited EBP reach under these conditions?***

Implementation Science (IS) - Determining what works, why and under what conditions

Background of the IS Problem: Understanding causes of limited EBP reach is critical to our stakeholders and to our field. IS seeks to determine why and under what conditions a strategy increases EBP reach.[93–95] Over 61 IS frameworks elaborate multiple domains,[96–99] and doing so, underscore limited knowledge of their underlying dynamic, multi-causal premises.[100] The Participatory System Dynamics (PSD) theory of change is that participatory stakeholder co-learning from modeling enables new capacities for managing system dynamics or mechanisms of change.[39,54] Informed by 60 years of PSD research, R01 specific aims define EBP reach as a system behavior, and test the PSD theory of change that grasp of system dynamics (mechanisms of change) is not just *for scientists*. Rather, systems thinking *by frontline staff,* is the basis for improving EBP reach. We briefly highlight limits of extant IS theories of change, and tests of mechanisms of change, focusing on what a PSD paradigm adds to IS, and its advantages over AF.

Theories of Change (TOC) are backward mapping methods for identifying the necessary determinants of change.[101] TOC is used as a planning strategy to link process and outcome.[102] These process frameworks are practical on face, but do not specify causal relationships that can be generalized.[32,103,104] For example, comparative qualitative analysis (CQA) is an inductive method to develop a TOC.[105,106] CQA retrospectively classifies sets of conditions co-observed with desired outcomes to descriptively infer causality, often identifying multiple pathways (permutations) to the same outcome (equifinality, see below).[105] Determinants

are not scientific theories[103] and pose risks as causal paradigms; Lack of tests for temporal precedence or effect size,[106] propensity for false positives[107] and confirmation bias,[108] all indicate caution. IS frameworks presume several components, but multiple comparisons inflate type I error,[109] impeding theory tests, just as finding multiple pathways leaves stakeholders weak prospective guidance for making EBP-related change.

Mechanisms of Change (MOC) are targets that will explain findings and accelerate research progress. Many implementation strategies produce no effect.[110] Review of multilevel mechanisms of implementation strategies in controlled trials, found nine tests of mediation, *all unsupported*.[111] Again, there are challenges; Confirmatory hypotheses require a strong theoretical and empirical base, as inclusion of multicollinear variables, and/or use of multiple tests both inflate error. Without a resource-intensive experimental design, use of mediating *associations* to infer causality is debated; 'Mediators differ from *mechanisms* which invoke a higher level of specificity and describe the precise sequence of operations or underlying causal processes through which an effect occurs' (Williams, 2016; p. 784). Finally, multilevel mediation is in the expert domain and inaccessible to most stakeholders, with limited utility for guiding local change. ***Is there an alternative?***

Preliminary Research – Two Causal PSD Theories for Improving EBP Reach – AF versus PSD (Aim 1)

Based on our R21 pilot, PSD is well suited to the need for improving EBP reach *and* need for IS progress in defining and testing TOC and MOC. The PSD paradigm sees the status quo 'limited EBP reach' as a function of the mental models that teams use to guide decisions, which are inadequate for redesigning EBP-related system dynamics.[36–38,47] PSD shows how 'today's problems come from yesterday's solutions.'[112] PSD simulation learning to make system causes transparent and analyzable, increases systems thinking enabling more effective ongoing change. PSD enlists two 'classic' causal theories: decision theory (Aim 2 - TOC) and systems theory (Aim 3 - MOC). Distinct from IS determinant frameworks and process models, 'Classic theories originate from fields external to IS, e.g., psychology…organizational theory, which can be applied to provide understanding and/or explanation of aspects of implementation' (Nilsen, 2015; p.8).[32,103,104]

AF is central to VA EBP implementation[15,50] and is one of the most common IS strategies used around the world.[29,30] AF and PSD each require data review.[34,89] But, PSD research challenges the sufficiency of data for selecting effective changes.[45,113] AF reviews have found variability in effectiveness could be traced to the complexity of the EBP (with AF more effective for simple practices[91]) and the feedback audience (with AF for team practices, more effectively delivered to teams[114,115]). AF must be frequent, in writing, and include a correct solution.[27,29] PSD agrees with need for timely, actionable feedback.[116,117] But, without systems thinking and modeling, PSD doubts the ability of policy makers to provide correct solutions for the myriad local issues that trouble EBP delivery.[118–120] PSD cautions, 'The cure can be worse than the disease.'[121,122] ***How might AF be counterproductive?***

> 'Despite being extensively studied, health care A&F interventions remain variably effective, with overall effect sizes that have not improved since 2003.'...'The development of the scientific basis of A&F in healthcare appears to have stagnated; we are not developing more effective A&F interventions than we were 20 years ago.'

PSD uses the term policy resistance[122] to describe the need for holism when making changes, represented in aphorisms such as 'The harder you push, the harder the system pushes back,'[112,123] and 'There are no side effects in systems. There are only effects.'[39,40,124] System resistance does *not refer to stakeholders' attitudes*; *it refers to powerful dynamics of the system*.

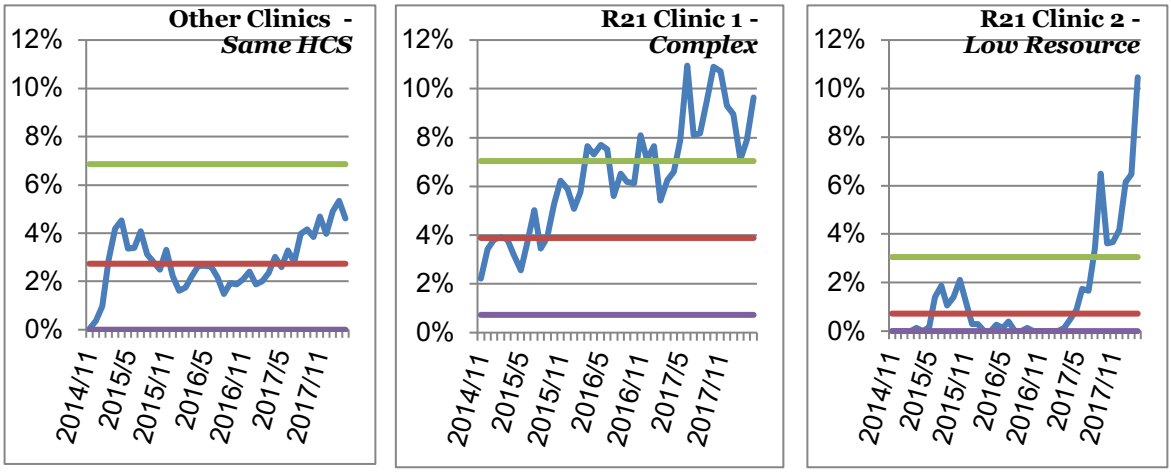
Preliminary Pilot. For example, prior to our R21 pilot study, our pilot partners invited us to try PSD after a previous SMART goal effort failed. The clinic set the following Specific, Measurable, Timebound goal: 'By April 2015, 40% of patients newly seen in outpatient mental health for depression, PTSD, or anxiety disorders will have two psychotherapy visits completed within 28 days from time of intake.' Staff saw initial improvement in scheduling (up from 25% to 65% scheduled), but wide unexplained scheduling variability was observed

over 9 months (some weeks 0% scheduled, some weeks 100%), and the gap between scheduling and completing 2 visits showed the 40% goal was never achieved (**new mean = 14%).** In SMART terms, PSD proposes that if the goal was never Achievable, because it was not Realistic in that Timeframe with the available resources, then it may do more harm than good by 1) generating further instability of the EBP-related system behaviors (see Aim 3 Background), and 2) undermining psychological safety, or willingness to learn and try out new solutions,[44,56,65] exacerbating staff burnout and risk for turnover (see Aim 2).[55,68,118]

Preliminary R21 Data in Support of Aim 1 – PSD Effectiveness

Our statistical process control analyses indicate our two R21 pilot clinics each demonstrated a three standard-deviation increase above their pre-intervention EBP reach using PSD (*Figure* 1; $\alpha < .003$). In *Figure 1*, purple = lower control limit; red = clinic 12-month pre-intervention EBP reach; green = upper control limit. Control limits are three standard deviations above and below the pre-intervention mean. These clinics have maintained this improvement for 12 months and 8 months respectively, whereas the other seven VA clinics from the same regional health care system (HCS) did not improve EBP reach over this period. Moreover, we observed no secular trend toward improved EBP reach in VA national AF measures over this period. These R21 quasi-experimental findings support the proposed R01 effectiveness test: a cluster randomized trial (CRT) to evaluate for PSD superiority over standard AF.

Figure 1. R21 Effectiveness – Statistical Process Control of EBP Reach Improvement (+3 *SD*)



Summary (Aim 1). AF and PSD research show data review is insufficient for effective change.[43,125] AF effectiveness is diminished[32] due to choosing ineffective strategies that leave system causes unaddressed.

Partnering to Define EBP Reach using *Systems Thinking* – PSD Theory of Change (TOC) – Aim 2

Why else doesn't AF work better? PSD explains that AF learning is also attenuated by delays between making changes and observing their real-world effects.[47] A recent scoping review of 'Healthcare Learning Organizations' found the majority principles outlined by System Dynamics (the field coined 'Learning Organization'),[121,123] including the central idea, in which systems thinking coheres mental models, shared vision, and team learning.[126] But, narrowed down from 263 keyword- identified articles, *only 2 published studies measured systems thinking*, the cornerstone of PSD research on change.

The PSD TOC defines learning as a system feedback process in which mental models are formed from feedback in the real or virtual world, which shape the rules used for decision-making, which then shapes the

real world. In other words, 'Seeing is believing *and* believing is seeing,'[39] especially as we act to change the real world. This learning process is more effective with PSD modeling, called 'double loop learning.'[124,127,128] From the first writings of the discipline,[129] PSD practices are guided by decision science.[46,130]

'Deep change in mental models, or *double loop learning*, arises when evidence not only alters our decisions within the context of existing frames, but also feeds back to alter our mental models. As our mental models change, we change the structure of our systems, creating different decision rules and new strategies. The same information, interpreted by a different model, now yields a different decision. Systems thinking is an iterative learning process in which we replace a reductionist, narrow, short-run, static view of the world with a holistic, broad, long-term,

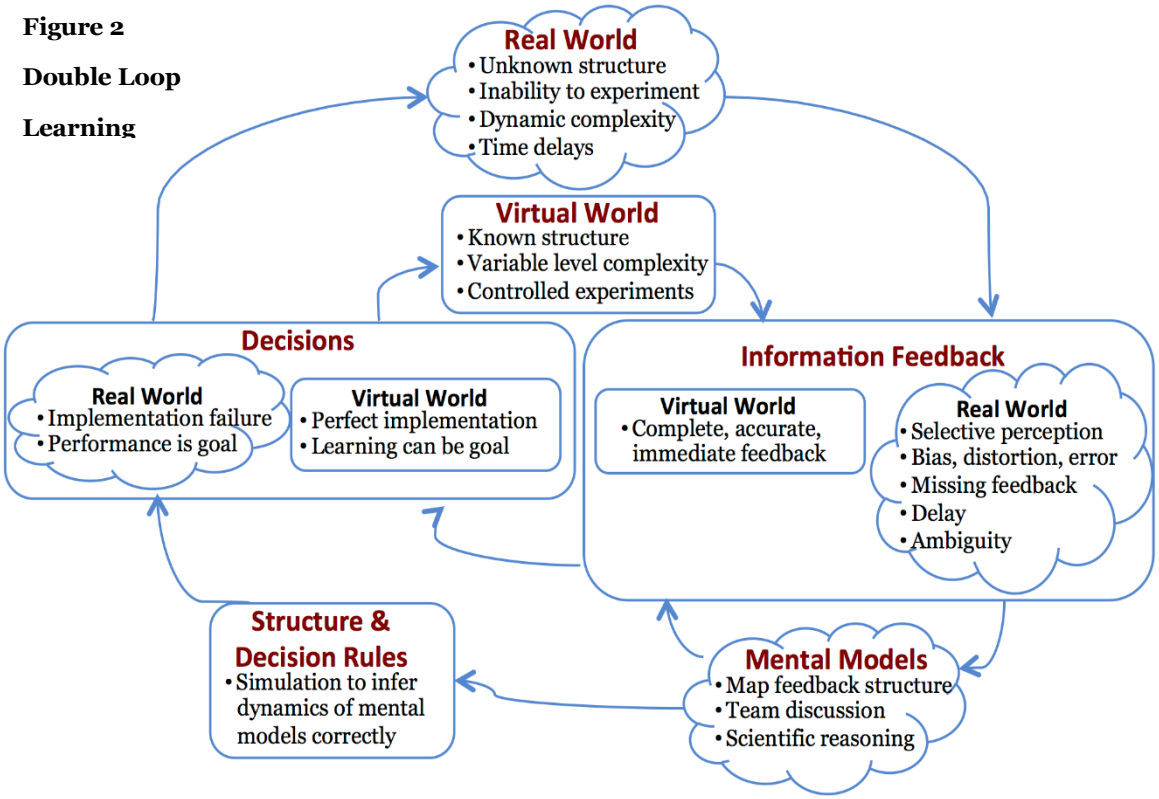
dynamic view, reinventing our policies and

PSD research identifies several cognitive biases and limitations that lead to poor decisions when facing complexity: use of heuristic mental models that seek minimally satisfying solutions rather than optimal solutions (bounded rationality),[130] rules of thumb that wrongly attribute the state of a system solely to inflows rather than outflows (correlational heuristic),[124] and inability to solve accumulation or delay problems (stock-flow failure).[113,124,131] A PSD insight is that in a complex system, cause and effect may not be closely related in time and space,[112] rendering learning from AF alone unlikely.

Systems Thinking as a Mediator of Change. *Figure 2* depicts the difficulty of learning from the complex, real world, as compared to the virtual world of modeling. PSD recommends simulation to improve mental models with systems thinking. Without PSD, defective causal decision rules impact explicit, effortful implementation planning (system 2 cognition) and implicit, automatic day-to-day decisions about EBP coordination and continuity (system 1 cognitions).[132–134] An effectiveness review of 107 PSD projects, identified more efficient improvements (33%), increased consensus (49%) and commitment to change (33%), including systems change guided and evaluated by modeling (42%).[43,125] A recent mediation study showed *increases in systems thinking due to PSD*, that led to increased psychological safety, which increased information sharing, explaining performance improvements.[44]

Figure 2

Double Loop Learning



Preliminary PSD Mixed-Methods Facilitation Data in Support of Aim 2 – Systems Thinking Mediator

Systems Thinking during PSD model development. We used PSD best practices, including nominal group

technique to generate 12 categories of 131 'limited EBP reach' related issues.[148,149] Over a period of six months of modeling we used several rounds of 'dot voting' to converge on PSD priorities: 'issues' were translated into 'variables' and 'decisions' in models.[136,150] Our purpose was to build models that explain 'limited EBP reach' *using providers' new systems thinking skills*. Our prompt question was: '*What simulation will help teams within the VA learn to manage the tradeoffs in how to provide evidence-based care to Veterans?'* Four Focus areas, 'Care Coordination,' 'Management Concerns,' 'Provider Capacities and Constraints,' and 'Provider Quality of Work Life' were narrowed to 9 specific priorities (*Table 2*).

Table 2. Staff Priority Experiments

- Patients' needs/preferences
- Reduce extra stops for Veterans
- Initiating a specific treatment
- Allocations of time (not enough time)
- Actual time (what we really do)
- Misunderstanding provider functions
- Morale & burnout
- Staff turnover

Summary (Aim 2). ***Systems thinking*** is the ability to recognize, understand, and synthesize interactions and interdependencies, including how actions and components can reinforce or counteract each other.[135]

Simulation learning upgrades mental models guiding daily system 1 (fast) *and* strategic system 2 (slow) EBP decisions in multidisciplinary teams, overcoming AF limits for learning *how* to improve local EBP reach[58,130]

Partnering to Define EBP Reach as a *System Behavior* – PSD Mechanisms of Change (MOC) – Aim 3

The ability to infer general principles from observations (generality) is foundational to scientific development. Use of PSD in Mayo Clinic identified substantial savings in the treatment of renal disease by recognizing that oscillating hemoglobin measures were *caused* by a mismatch between the measurement and its use in guiding clinician medication decisions.[151,152] Although hired to develop an AF system, use of PSD with a multidisciplinary team recognized the problem was across hematology and nephrology. Based on the measures available, clinicians were administering a second dose before the first took effect, causing clinically acute adverse effects. *Identification of this underlying biophysical dynamic was generalizable and led to Mayo Clinic-wide implementation of PSD to guide individualized dosing*, which brought population-level hemoglobin within range, increased well-being of patients, and reduced costs.[151,152]

Aim 3 will test whether R21 models have potential to replicate this type of advance in addiction and mental health. We hypothesize the structural-behavioral validity of our R21 models will generalize across AF/PSD arms.

In common with *Community Based System Dynamics*[136] we locate PSD within the continuum of participatory research.[48,137] PSD is a partnership approach that equitably involves all stakeholders' expertise, in all aspects of the research development process, using shared decision-making activities that are designed to produce system change.[137–140] We committed to equitable PSD resource development, valuing local staff knowledge in PSD models and activities. Unlike most AF systems, we co-created our new shared PSD assets with frontline staff.[48] **PSD activities elicited stakeholders' *mental models* about how they think EBP implementation works, with those interconnections made explicit in PSD models.**

A system is set of elements interconnected in such a way that they produce their own internal dynamics.[122] The dynamics of a system problem *cause* its behavior. Many IS frameworks, and stakeholders, view a system as an external setting or organizational context (exogenous).[98,141–145] PSD does not. EBP reach behavior emerges from internal causes.[122,124] Using AF, these system dynamics are hidden 'black boxes.' With simulation, PSD makes causal dynamics transparent in real-time. The PSD endogenous theory of 'limited EBP reach' is qualitatively refined with stakeholders and rigorously assessed for structural-behavioral validity.[49,51,146,147] The PSD endogenous view is empowering, proposing that local teams engage in mutual learning to co-create solutions that change the dynamics of EBP reach.

PSD Model Structural-Behavioral Validity is present when the model represents its purpose defining 'limited EBP reach,' and the accuracy of its formulation is rigorously confirmed using reliable data. [49,51,124,147,153] All of these conditions were present in our R21 model validation using VA data.[15,48,50] Tests of structural-behavioral validity show how PSD reconciles the equifinal/multifinal IS paradox. The equifinal and multifinal columns in *Table 3* display explanatory causal operators (structure of the equation), accounting for specific numerical values to derive the result.

Table 3. Causality

Equifinal	Multifinal
1 + 1 = 2	2 x 3 = 6
1 - 1 = 0	3 x 4 = 12
1 x 1 = 1	5 x 1 = 5
1 / 1 = 1	0 x 6 = 0

In the equifinal column of *Table 3*, mathematical operators demonstrate *different* causal relationships sometimes achieve the *same* result (2 out of 4 times), and sometimes don't, even with the *same* numerical values. The multifinal column demonstrates the *same* cause achieving *different* results as a function of *different* numerical values. These mathematical facts demonstrate the key import of understanding causes.

PSD formulates generalizable feedback and stocks-and-flow equations that produce EBP reach according to system theory, when parameterized with local data to guide change. The fundamental insight of PSD*, 'Dynamics (equations) before details (specific parameter values),'* is drawn from the fundamental theorem of calculus.

Preliminary PSD Structural-Behavioral Validity Data in Support of Aim 3.

Causal Stock and Flow Dynamics. We found that although stakeholders may disagree about details, rarely did they disagree about dynamics. For example, several insights are visual *and* clear in *Figure 3*, which shows how changing the 'Medication Management (MM) Appointment Frequency' and 'MM Return Visit Interval' are related, *and* influence the interdependent 'New and Existing Patient Services,' *and* how the units of appointments ('MM Appts scheduled'), 'MM New Patient' start rate and 'MM Patients waiting for 1st visit' are all related. Before calibration, these interdependencies were refined qualitatively to 'saturation.' After several iterations with multiple teams and stakeholders no new key dynamics were identified. These issues are generic across a variety of clinical teams.

Table 4. VA Corporate Data Warehouse (CDW) Source for Model Inputs

Patient Cohorts	ICD diagnostic information from encounters with providers
Clinic Hours	Clinic scheduling hours & availability from VA scheduling system
Provider Specialties	Provider disciplines with user input to map to encounter types
Clinic Appointments	Completed CPT encounters. Appointment data = 'no shows'
EBP reach	EBPsy [illegible]plate usage or VA SAIL [illegible]ata

ICD = International Classification of Disorders. CPT = common procedural terminology. EBPsy = Evidence-based psychotherapy. EBPharm = Evidence-based pharmacotherapy. SAIL = Strategic Analytics for Improvement Learning (quality measures, see also *Table 1*).

Improving team understanding of these dynamics was deemed critical by staff, particularly with regard to OUD EBPharm, which requires a very strict return, visit interval for an adequate dose among existing patients. This *tradeoff* with 'Decision to start Patient in MM' is typically managed with mental heuristics, which staff agreed would be better optimized with PSD modeling. NOTE: *Figure 3* is static, but PSD simulations are temporally dynamic and show the impact of decisions in 1) real-time, on 2) EBP reach

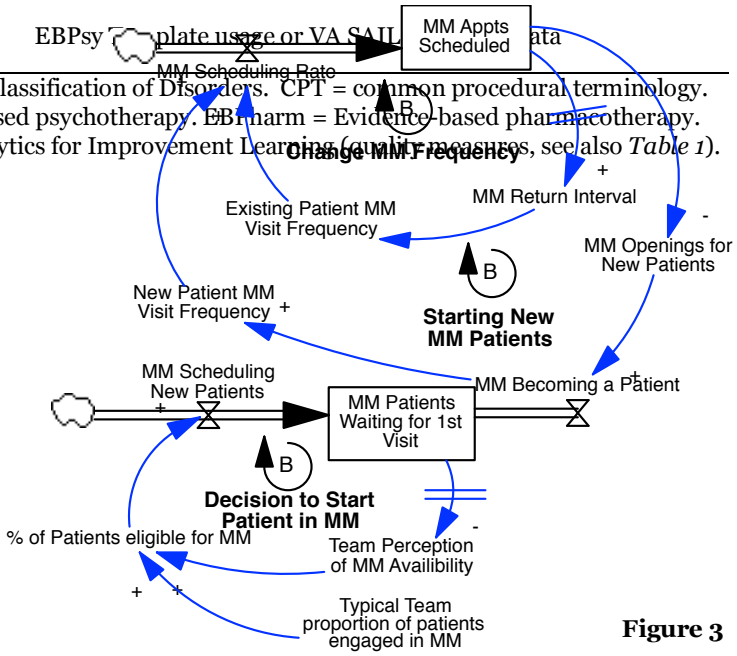


Figure 3

over a variety of future time horizons. We noted intent to submit a video for the R01 to show how these causal dynamics are displayed in the 'Modeling to Learn' Team Training (described below).

Calibration and Validation of PSD Models. PSD models are calibrated against existing EBP reach time-series data for each decision-making team, using Kalman Filtering and Monte Carlo methods for optimization.[49] Sensitivity analyses test the structural-behavioral validity of the model, which also was formulated to be consistent with seminal service system PSD models.[154]

Figure 3 and *Table 4 and Table 5* (on the next page) show PSD model dynamics and the associated parameter inputs are expected to be generic across a variety of services, patient populations **and EBPs.** But, parameter *values* are team-specific to guide local decision. *Like the Mayo Clinic example, this enables participatory 'modeling to learn' activities at scale. Table 5* displays example model inputs, feedback dynamics represented as provider decisions, and model equations. In addition, based on feasibility and acceptability of PSD in the R21, **NCPTSD funded the development of 'Modeling to Learn.'** This workshop and user-interface enables this specification and sophistication under the hood, and real-time causal learning simulations to frontline EBP decision-makers. NOTE: Demonstration of real-time dynamic interface used in team training is available in the 2-minute 'Modeling to Learn' simulation video.

Summary (Aim 3). ***Generalizability*** PSD effectiveness of improving reach in a variety of settings is enhanced by the use of standard data definitions for model inputs and the generality of using mathematical principles to identify the common dynamics of EBP services systems. Based on the multiple iterations of stakeholder engagement and structural-behavioral validity testing, we expect that our four PSD models will explain reach in any R01 study clinic.

Summary of Preliminary Data in Support of R01 Specific Aims

We propose a Phase III Clinical Superiority design with falsifiable hypothesis tests (see *Table 6*). We propose the two-arm CRT to test theoretical, *confirmatory* effectiveness (Aim 1), causality (Aim 2) and generality/generalizability (Aim 3) hypotheses. The proposed R01 is responsive to calls to develop AF strategies that are theoretically rigorous due to ineffectiveness of AF to date. We test AF (described below), against very rigorous theory-based approach (PSD) with a 60-year track record of effectiveness for improving organizations. Building from the PSD research tradition, proposal to conduct robust tests of internal validity and external validity is relatively rare in the IS field, which is often exploratory, especially regarding causality. We focus on EBP reach, due to solid EBPsy/EBPharm Effectiveness, and hefty VA Adoption, Implementation and Maintenance efforts (Re-AIM).[93] We locate R21 insights in implementation science (IS), outlining our rationale for expected PSD superiority (Aim 1), based on PSD theory of change (Aim 2), and PSD mechanisms of change (Aim 3).

Table 5. Model Inputs, Feedback/Decisions & Equations (example)

Inputs (care coordination model)

New Patients Per Week	
One and Done Patients (%)	Mean Usage (appointment/week)
Missed Opportunities Rate (no show)	Median Usage (appointment/week)
Mean Return Visit Interval (weeks)	Seventy Five Percent Usage (appt/wk)
Median Return Visit Interval (weeks)	Mean CC Engagement (duration)
Mean Wait to First Visit (weeks)	Median CC Engagement (duration)
Median Wait to First Visit (weeks)	

Feedback & Decisions (medication management model)

Focus on Actively Engaged MM Patients, Work more MM hours, Change MM Frequency, Starting New MM Patients, Decision to Start Patient in MM

Model Equations (aggregate model)

Ideal Patient Start Rate for Available Hours[service]=(Ideal Completion Rate for Available Hours[service]/Demand per Patient per Week[service])/Actual Time in Service[service]; **Units** = patient/Week

Maximum Start Rate[service]=Patients Waiting to Start Service[service]/Minimum Time to Schedule[service]; **Units** = patient/Week

Ideal Completion Rate for Available Hours[service]=Baseline Service Capacity[service]/Time per Appointment[service]; **Units** = appt/Week

Table 6. Summary of Proposed R01 AF/PSD CRT Design, Hypotheses, Tests & Measures for Specific Aims

Aim	Purpose	Hypothesis	Test	Measure
1	**Effectiveness**	PSD will be superior to AF for improving EBP reach	Cluster Randomization	Existing Patient Data; 6 & 12 month EBP initiation and course
2	**Causality**	Effect of PSD/AF on EBP Reach explained by Systems Thinking	Multilevel Mediation	6 mo. Team STS Mean (Mediator); Existing Patient Data (Outcome)
3	**Generality**	SD models will explain EBP reach across clinics in PSD AF arms	Structural-Behavioral	Match of CC, MM, Psy, Agg Model Structures to Observed EBP reach

RESEARCH DESIGN AND METHODS

Overview. Five standards guided study design and will be followed for reporting (*Table 7*).

Table 7. Design and Reporting Standards

Cluster Randomized Trial Design	**CONSORT***
PSD Simulation Model Documentation	**SDM-DOC**
PSD Simulation Model Output	**SIMULATE**
PSD/AF EBP Education Intervention	**GREET**
MH/Addiction Common Data Elements	**PHENX**

* 2010 update for 'parallel group randomized' trials. MH = Mental Health.

Protocols for AF and PSD Implementation:

Protocols were developed to be implemented and evaluated for fidelity (see *Table* 9) using the Guidelines for Reporting Evidence-based Practice Educational interventions and Teaching (GREET). Led by the R21/R01 team (*Table* 8), National Center for PTSD (NCPTSD), Office of Mental Health and Suicide Prevention (OMHSP), Office of Strategic Integration (OSI) and Veterans Engineering Resource Center (VERC) developed a website for reviewing team data (AF/PSD), and a second website for team simulation (PSD). AF *'Team Feedback'* clinics will be trained to use the data tools in 2 team huddles (month 1), with weekly emailed data reports linking to the data (months 2-6). PSD *'Modeling to Learn'* clinics will use the data website, simulation website, and sessions with Care Coordination (CC), Medication Management (MM), Psychotherapy (Psy), and Aggregation (Agg) models, ending with facilitated wrap-up reflection (months 2-6).

AF Dynamic Data Tools. The AF condition is comprised of tools for reviewing data before synthesis in PSD simulation models is graphically displayed as retrospective trends overtime for the team. Trends graphs also produce searchable reports. The team can review their typical screening practices to note where gaps in quality exist. They can also review the underlying report to follow-up with specific patients who require care coordination and follow-up. Teams value graphical review of EBP template data.

PSD Simulation Tools. Data review improves provider confidence in the data that feeds the simulation models. Augmenting AF data review, PSD model simulations are designed to more precisely guide decisions through exploration of the expected causes of current EBP reach and simulations of the impacts of changes in future-oriented projections or scenarios.

Team Training Sessions. Virtually facilitated AF/PSD workshop sessions (see *Table* 9) are designed for teams of providers (typically 5-10/team) from each clinic, with no cap on total providers (typically ~20-40 staff). Frontline leadership and one 'champion' from each service delivery team will receive additional PSD resources to operate as an internal facilitator for their team. Mental health staff will be eligible to receive continuing education credits for AF (2 credit hours) and PSD (12 credit hours) workshops from VA Employee Education Service (EES) for six primary frontline disciplines: psychiatry, psychology, social work, counseling, nursing and certified peer-support.

Table 9. 'Team Feedback' (AF) and 'Modeling to Learn' (PSD) Protocol Implementation and Fidelity

Arm	Resources	Session	Use	Objective
AF & PSD	*EES Virtual Workshop Team Data Sharepoint	Intro (1a) Data (1b)	# Learners # Unique Visits	Increase knowledge of current EBP reach
AF	Weekly Email w/Data Report	Self-directed (2-6)	# Clicks Min. on Site	Identify EBP reach gaps for improvement
PSD	Team Data Sharepoint Simulation Interface Weekly Email Data & Sim. Report	CC (2) MM (3) Psy (4) Agg (5) Wrap (6)	# Clicks Min. on Site(s) # Simulations # Sims Shared Sim. Inputs & Outputs	Apply systems thinking & simulation to see causes of EBP reach & run tests to find improvements

*EES – Employee Education System accreditation for physicians, psychologists, social workers, nurses, counselors and certified peer specialists (AF = 2 credit hours; PSD = 12 credit hours;). Min. = Minutes. Sim = Simulation. Models: CC = Care Coordination. MM = Medication Management. Psy = Psychotherapy. Agg = Aggregate.

Project Management Plan

Study Team and Stakeholders.

See *Table 8* in Section 3.5 – Structure of the Study Team – Human Subjects and Clinical Trials.

Study Timeline and Feasibility

See *Table 10* in Section 2.7 – Structure of the Study Team – Human Subjects and Clinical Trials.

We propose 60 total months of study activities and 30 months of active PSD facilitation or post-training technical assistance (phases 3 and 4). This leaves 24 months of flexibility for delays across pre (phases 1 and 2) and post (phases 5 thru 8) activities (*Table 10*).

Outcomes and Analysis Plan – Definitions and VA Corporate Data Warehouse (CDW) Data Use

Evidence-based Practices (EBPs). We selected EBPs for highly prevalent AUD, OUD, PTSD and depression[4] based on demonstrated clinical efficacy and effectiveness[76–82,87,88] and limited EBP reach (*Table 1*).

EBP Reach definition. We propose to use PSD to improve reach of 7 EBPs in the outpatient system. We define reach as the proportion of patients diagnosed with OUD, AUD, PTSD, or depression (ICD-10 codes) who meet EBPsy and EBPharm *1a) initiation* and *1b) course* measures (numerator) divided by the total number of patients with these diagnoses (denominator) at that clinic. Initiation of an EBP is indicated by EBPsy template or EBPharm prescription after intake. Adequate course is based on receiving an adequate number of EBPsy sessions to be a "completer" (typically 8 sessions) or enough refills for a guideline-recommended adequate trial of each medication (varies by medication) (see CDW/SAIL detail in *Table 1)*.

EBPsy Definition. We will study five **EBPsy.** Three for depression (Cognitive Behavior Therapy [CBT-D], Acceptance and Commitment Therapy [ACT], Interpersonal Psychotherapy [IPT]), and two for PTSD (Prolonged Exposure [PE], Cognitive Processing Therapy [CPT]).

EBPsy Measure. Our EBPsy measure is completion of EBP templates during sessions with a relevant CPT code.[156] EBP templates may not reflect overall EBP delivery without templates. However, the templates confer several significant advantages, including 1) national standardization, 2) maximizing use of existing CDW data for AF, 3) integration with SAIL quality measures (PTSD 56) and EBP dissemination programs.

EBPharm Definition. Our EBPharm measures follows VA SAIL definitions (SUD 16, MDD43h, MDD47h): a combination of prescriptions placed with the VA pharmacy and sessions with a relevant CPT code.

EBPharm Measure. We will also improve reach of eight EBPharm measures - two for depression (84 and 180 days therapeutic continuity at new antidepressant start), two for OUD (methadone and buprenorphine) and four for AUD (Acamprosate, Disulfiram, Naltrexone, and Topiramate).

Team Survey Measures

Participatory Measure Selection and Evaluation Process. AF and PSD theories of change and prior empirical research led us to target measures of: 1) systems thinking, 2) learning, 3) team decision-making, 4) psychological safety, 5) burnout/morale, and 6) pragmatic measures of PSD/AF implementation. We reviewed Society for Implementation Research Collaboration and Agency for Healthcare Research and Quality Measure Repositories, measures in *Implementation Science* and *System Dynamics Review,* and key IS review manuscripts,[157–159] noting measures validated with the VA staff. After identifying the most valid and reliable measures (see *Table 11*), our team of patients, providers, manager and program leads further reached consensus about the measures based on face and content validity, or factor structure. Higher weight was given to brief measures and measures preferred by providers, reaching a final survey items to be completed by staff pre/post AF/PSD.

Measurement Schedule. *Table 11* shows the plan to collect 3 measures at baseline (BL) and 6 months (52 items), our mediator (STS) and measures to describe team (PACT) and clinic (LOS-27) contexts. Two measures of AF/PSD practicality (AIM, IAM, FIM), and AF/PSD effect on team decision-making (TDMQ) are added at 6 months (83 items).

Table 11. Team Measures

Scale	Aim	Time	Items	Valid	Reliable	Factors
1. Systems Thinking Scale (STS)	***Aim 2***	BL & 6 mo.	20	0.89	0.74	1
2. Patient Aligned Care Team (PACT) Survey-Burnout	Desc.	BL & 6 mo.	5	0.89	n/a	n/a
3. Learning Organization Survey (LOS-27)	Desc.	BL & 6 mo.	27	0.75 to 0.93	0.67 to 0.92	7
4. Team Decision Making Questionnaire (TDMQ)	Desc.	6 mo.	19	0.96	0.52 to 0.94	4
5. Pragmatic Measures (AIM/IAM/FIM)	Desc.	6 mo.	12	0.85 to 0.91	0.73 to 0.88	3

Systems Thinking Scale (STS Mediator). We selected the STS as mediator measure due to demonstration of criterion-validity. In addition to the measure details reported in the STS manual, the PI spoke to the authors to confirm appropriateness for use testing PSD theory of change in this R01 (Aim 2). The STS was developed as part of the Advance the Science of Quality Improvement Research and Evaluation (ASQUIRE) initiative and validated among health care staff for purposes of quality improvement, including evidence the STS was manipulated by systems thinking education (sensitivity to change)[160,161] *and* STS scores discriminated 3-groups based on their level of systems thinking education (Cohen's effect size d = .78).[135] The authors used the Quality Improvement Knowledge Application Tool (QIKAT),[162] as a measure of convergent validity. But, the relatively low STS/QIKAT correlation (.28-.46) provides discriminant validity for our purposes. Based on this finding, the STS should distinguish general quality improvement knowledge from AF) from systems

thinking (from PSD). After dimension reduction, the 20-item single-factor STS has strong internal consistently (.89) and test-retest reliability (.74).

Team Decision Making Questionnaire. The TDMQ asks, '*To what extent does AF/PSD help you to…*' This four factor scale was validated to assess the impact of a team intervention on team decision-making, support, learning and development of quality services.[66]

Exploratory Aim. Context Measures for Description of Study Clinics and Posthoc Analyses

Patient Aligned Care Team - (PACT). From VA team-based primary care, this 4-item descriptive measure tracks 1) years of experience with the team, 2) working on more than one team, 3) turnover/change in team staff, 4) team overwork, and the single-item 5) self-reported burnout (sensitivity 83.2 % and specificity 87.4 %)[67–70,118]

Learning Organization Survey (LOS-27). Developed out of the learning organization tradition,[65] the 27-item Learning Organization Survey demonstrated good psychometric properties during VA validation,[65] and we include it to assess 7 clinic context factors: a) supportive learning environment (including psychological safety), b) leadership that reinforces learning, c) experimentation, d) training, e) knowledge acquisition, f) time for reflection, and g) performance monitoring. Response scales range from 0 (never) to 4 (always) for subscales a, and c-g. Subscale b ranges from 0 (highly accurate) to 7 (highly inaccurate). Items will be reverse scored as necessary and summed for analyses (in *Appendix)*.

Measures of Feasibility (FIM), Appropriateness (IAM) and Acceptability (AIM) will assess for differences in team perceptions of PSD and AF on these three factors (12-items). Despite similarities, AF and PSD may be perceived differently in pragmatic terms. These scales have strong psychometric properties (*see Table 11*) for implementation research. Use in the R01 contributes to replication of reliability and dimensionality, and public release of data from the proposed sample of 720 providers adds to future measure norming.

Systems Thinking Mediator: Multilevel Psychometric Analytic Plan

Baseline Measures and Descriptive Analyses. STS means, standard deviations, and correlations will be calculated for individuals and teams. Intraclass correlation coefficients (ICCs) will be calculated. ICC values represent team 'traits,' estimated as the proportion of total measure variance attributed to variance between teams. An ICC near zero indicates highly variable STS within teams. We will do the same calculation for clinics within regional VA health systems.

Multilevel Confirmatory Factor Analysis (MCFA). Factor analyses will be confirmatory. Psychometrics and factor structure for the STS was determined in prior research (*Table 11*). Our unit of interest is teams. Teams are the unit for EBP implementation and PSD/AF training. As a result, we will use the STS team-mean. MCFA will be used due to nested measurement (teams in clinics), and potential for differential measure performance across arms in Aim 2 analyses due to PSD/AF assignment. MCFA with R package 'lavaan,'[163] will determine whether the same covariance (factor) structure holds for clinics in each arm (PSD/AF) via closeness of fit (CFI), root mean squared error (RMSEA) and information criteria (BIC/AIC). MCFA reduces risk for inflated error variance associated with non-independent measures (teams in clinics), which is known to inflate type I errors in association tests (mediation). MCFA ensures valid reliability and validity estimates of theorized constructs, and ensures valid tests of Aim 2 mediation hypotheses. The single-factor structure will be specified consistent with prior STS psychometric validation and results reported.

Criterion/Predictive Validity. We expect systems thinking to be sensitive to change and discriminate between our PSD/AF arms at 6 and 12 months. We will conduct a cross-arm manipulation check of PSD/AF assignment on systems thinking. Measures will be standardized and baseline assignment will be assessed

for an increase in systems thinking among PSD teams. We expect an effect size consistent with STS validation research showing a different between groups based on receipt of systems thinking training (*d* = *.78*).[135]

Reliability. Pre/Post Measurement Invariance will be examined with generalizability coefficients (GC)[164] in R package 'lavaan,'[163] which extend tests of internal consistency to designs with multiple sources of error. GC values measure the true variability of STS team-means in each arm as a proportion of total variance to determine the reliability of the scales across arms over time (Factor loadings and 95% confidence intervals). A standard set of increasingly constrained SEM Models will assess factor loadings, intercepts and residual variance for equivalence[165] and will detect the magnitude of measure non-invariance to obtain correct mediation inferences. Variance- covariance matrices of within- and between-arm latent factors will be plotted.

Qualitative and Observational Measures

Systems Thinking Codebook and Session Observations. Workshop Coding**.** We selected the STS self-report measure due to validation among health care staff receiving quality improvement training.[135] However, the primary systems thinking measures in *System Dynamics Review* are observation of systems thinking in language/explanations, and performance demonstrating system thinking skills (competence). Potential for coding bias precluded use of coding measures in mediation analyses, compared to the effect size estimate available for the STS. However, observational coding of online AF and PSD sessions, confers strengths for describing use or non-use of systems thinking to specific EBP-related decisions across AF/PSD arms.

AF/STS Fidelity and Use. Workshop Coding. We will review fidelity with qualitative checks against AF/PSD facilitator scripts for session learning objectives, 'key idea' and 'definitions,' including tracking the proportion of AF/PSD session activities (in minutes) on these components (see *Table 9*). Web-based Observation. Use of online AF and PSD resources will be tracked for criteria listed in *Table 9* across the active AF/PSD phases, and for ongoing AF/PSD sustainment (see Timeline *Table 10).*

Design for Tests of Specific Aims

We propose a randomized, two-arm, parallel group CRT to test theoretical, *confirmatory* effectiveness (Aim 1), causality (Aim 2) and generality (Aim 3) hypotheses. To avoid inflating type I error for Aims 1 and 2 we will conduct two omnibus tests for improved EBP reach (initiate and course). Due to monthly fluctuations in EBP reach, we will calculate 12-month period pre/post EBP reach averages, which removes clustering of EBP reach observations *within* clinics over time in tests of specific aims.

See *Table 10* in Section 4.2 – Statistical Design and Power – Human Subjects and Clinical Trials for detailed information on our calculation of Intraclass Correlation Coefficient (ICC), Average Clinic Cluster Size, Power Analyses and plans for Attrition.

Analyses for Tests of Specific Aims

***Aim 1*. Test for superiority of PSD over AF for increasing EBP initiation and course.** We will use R for tests of Aims 1a and 1b to establish PSD superiority. First, we will assess ICCs for within and between clinics. We will test for a significant difference between PSD and AF arms in increasing EBP reach using two (initiation and course) generalized estimating equation analyses for differences in proportions (reach).[57],We will estimate two (initiation and course) generalized linear models that account for clustering. We will assess distribution of EBP reach to identify the appropriate link function for robust standard errors, and test for a significant difference in EBP reach between the PSD and AF arms. We will report the effect size and 95%

confidence intervals for the difference between arms at alpha = .05.[57]

***Aim 2*. Test the PSD theory of change that increased EBP reach is via systems thinking.**[47,58,59]

We hypothesize that the effect of PSD/AF

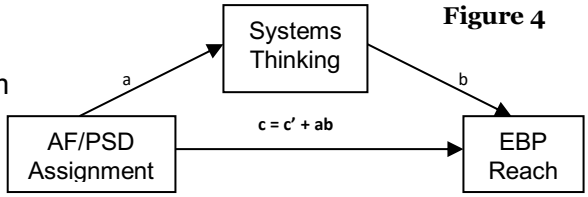


Figure 4

on 12 month period EBP reach will be explained by 6 month team systems thinking (STS), adjusting for baseline covariates using the ratio of mediator probability weight (RMPW).[60–62] Our CRT design uses AF/PSD assignment (independent variable) to experimentally manipulate STS (mediator) on clinic-level EBP reach (outcome) in each arm. But, STS is not randomized. RMPW uses sensitivity tests to address potential bias due to an interaction between the intervention and the mediator, or due to the operation of systems thinking through unhypothesized mechanisms (confounders) using multilevel mediation analyses in R package 'MultisiteMediation'[62] a confirmatory hypothesis test using multilevel mediation with partial variance in 6 month STS and 12 month period EBP reach across arms. *Figure 4* displays the mediation model from strategies (baseline), to mediators (Baseline to 6 month STS change), to outcomes (12 months), where ***a*** = the standardized beta coefficient of PSD/AF assignment on systems thinking, ***b*** = the effect of systems thinking on EBP reach, ***c*** = the total effect of PSD/AF assignment on EBP reach, ***c'*** = the direct effect of PSD/AF on EBP reach, and ***ab*** = the indirect effect of PSD/AF on EBP reach through STS (hypothesized mediation). We will use a bootstrapping, asymmetric confidence interval approach to balance power and type I error.

***Aim 3*. Test the generality of mechanisms of change in EBPsy/EBPharm SD models.** We will use structural-behavioral validation tests[51,63,64] to evaluate whether causal mechanisms formulated in our Psy/EBPsy and MM/EBPharm models generalize to explain EBP reach as a function of local data across PSD and AF clinics. *Figure 5* displays the analytic plan we will use to establish the Structural-Behavioral validity of the two models against observed 12 month EBP Reach in both arms (AF and PSD), using the Theil Statistics module in Vensim DSS[49,170]

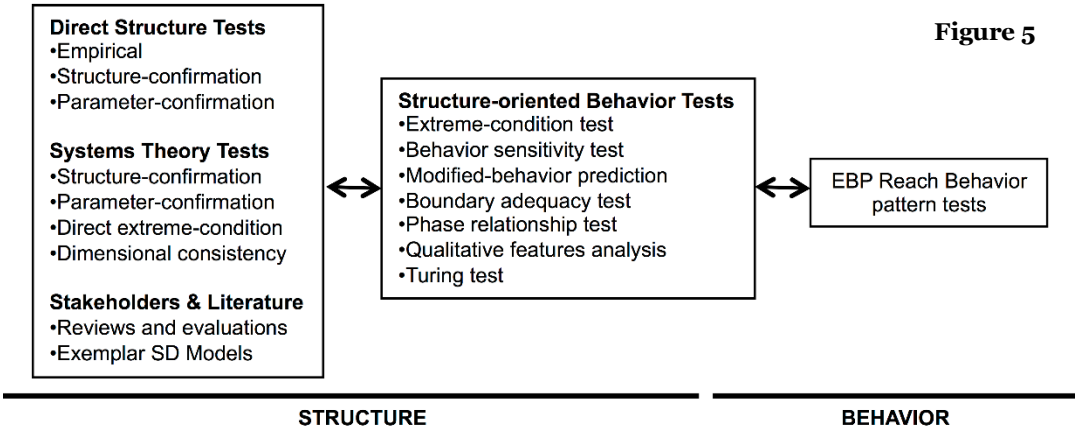


Figure 5

PSD Models. We will document our models using SDM-DOC and report findings using SIMULATE.[49,50]

Limitations and Alternatives

Limitations. Random assignment to usual AF or PSD is the best comparator for guiding decisions about scaling/sustaining PSD infrastructure. However, clinics in our 'usual AF' comparator will likely engage in highly variable AF activities. We take steps to address this with CDW-based stratification.

Alternatives. We will complement our primary tests of Aim 1a and 1b using autoregressive integrated moving average (ARIMA) to mitigate potential limitations.[171,172] This helps to guard against threats from secular trends (overall patient demand or VA EBP adoption may be increasing), and regression to the mean (given these are lower performing clinics). We will track national trends in EBP reach and estimate the percentage of

regression towards the mean.[173,174] We will also graphically display 12 month period averages of EBP reach related to Aim 1a and 1b at the clinic level, using statistical process control (SPC). SPC p-charts will display the pre/post intervention EBP reach proportions for all PSD and AF clinics. SPC is a standard, healthcare quality tool robust for non-normal data and unbalanced samples.[175] The p-chart centerline corresponds to the mean proportion of patients who meet EBP criteria, controlling for the number of patients in each observation.[176,177] We will use SQUIRE standards for reporting SPC.[177,178]

INNOVATION

Generalizable theoretical and empirical mechanisms to empower everyday EBP decision-makers. This proposal tests the effectiveness, causes and generality of PSD against a highly prevalent comparator in the largest integrated health network in the U.S. The PSD approach empowers frontline teams to directly test their own local causal attributions for EBP reach, using theory-guided, mathematically specified simulations and existing health system data from their own team.[38,85] Due to NCPTSD development of 'Modeling to Learn,' PSD learning can be scaled with online resources and virtual facilitation.[39,44]

PSD is innovative meeting several needs for advancing the science of EBP implementation (*Table 12*). Health care data systems are increasingly standardized (e.g., ICD and CPT codes).[179] Lessons learned in VA outpatient addiction and mental health can be translated to other VA health services,[180] and health care systems.[181] The public health care and education missions of NCPTSD enable dissemination of free, fully transparent, open source PSD facilitation scripts, PSD models, and SQL code examples for standard EHR data extraction. The proposed CRT focuses on the significant problem of improving EBP reach, and advances IS with a paradigm that has potential for a significant pay-off. Due to the use of high-quality tests, *null findings for Aims 1-3* will also inform the IS discipline increasingly turning to operations and systems science.[103,182–188] Our ability to formulate robust hypothesis tests is based on the translation of analogous fields to IS, rigorous qualitative and quantitative R21 preliminary research, and our ability to capitalize on the VA AF infrastructure, enabling a powerful comparison of PSD against one of the most commonly used strategies in the world.

Table 12. PSD Insights for Stakeholders & Implementation Science

1. Range of system behaviors explained by causal dynamics
2. Generalizable variable relationships across settings; Standard data inputs
3. Primary role of dynamics in producing EBP behavior
4. Secondary role of specific variable values in producing EBP behavior
5. Qualitative co-development and mutual learning with stakeholders
6. Quantitative testing of causal assumptions (feedback dynamics)
7. Causal mechanisms matched to local capacities/constraints (local data)
8. Intuitive visuals & 'system stories' for frontline teams (EBP implementers)
9. Without PSD underlying causal EBP dynamics is not obvious (black box)
10. Simulation an efficient, highly-effective way to select implementation plans
11. Experiential learning improves day-to-day EBP-related decision making

STATISTICAL DESIGN AND POWER

DESIGN

We propose a parallel two-arm, 24-site (12 clinics/arm) cluster randomized trial (CRT) to establish superiority of PSD over usual AF for increasing EBP reach. CRTs are best for complex interventions like PSD, with many interacting components, and in which the unit of intervention and observation is the clinic.[57] Number of clinics and number of patients per clinic define total CRT size. Our primary aim is to increase the proportion of the patient-population within PSD clinics that receive EBPs (reach: initiate/course). Cluster size is defined by the eligible depression, PTSD, OUD patient cohorts. Clinic proportion (reach), is the sum of patient-level EBP reach in usual care: patients receive (1) or do not receive (0) an EBP (binary).

Inclusion/Exclusion Criteria for Study Clinics. We will randomize and complete the PSD '*Modeling to Learn'* virtual workshop with VA divisions and community-based outpatient clinics (CBOCs) or 'clinics.' Eligible clinics will be from regional VA health care systems (HCS) below the overall SAIL quality median, and 3 of 8 SAIL measures associated with 4 EBPsy and 3 EBPharm for depression, PTSD, and OUD in *Table 1*. Inclusion criteria balance sensitivity and specificity in identifying clinics from lower performing HCS.

Once identified via SAIL/CDW, recruitment will occur via OMHSP, NCPTSD, EBP Coordinator, and VISN Mental Health Lead networks (see Section 2.5 - Recruitment and Retention Plan). Clinics must have HCS director assent to randomization. Each clinic will be from a separate health care system (1 clinic/health care system). Analyses will control for clinics nested within VA.[57] Given known PSD interest, and R01 partner support, recruiting 12 clinics from these networks should be highly feasible, enhanced by 18 months total pre/post flexibility (see *Table 10*).

Stratification. Using baseline CDW data, we will conduct computer-assisted stratified block randomization with the R package 'blockrand'[155] to balance arms for region, and 5 factors expected to influence clinic-level EBP reach: baseline EBP reach, clinic size (defined by two measures, both patients and providers), urban vs. rural location[3] and medical division vs. community-based outpatient clinic (CBOC). We stagger start dates every 6 months (3 PSD clinics/3 AF clinics per wave), to mitigate cohort effects, making management of multi-site relationships feasible.

EBP Pre/Post Operationalization. For each wave, we use 12-month period average of EBP reach before AF/PSD start (pre-measure) and 12-month period average of EBP reach after AF/PSD end (post-measure).

ANALYSES FOR TESTS OF SPECIFIC AIMS

We propose a randomized, two-arm, parallel group CRT to test theoretical, *confirmatory* effectiveness (Aim 1), causality (Aim 2) and generality (Aim 3) hypotheses. To avoid inflating type I error for Aims 1 and

2 we will conduct two omnibus tests for improved EBP reach (initiate and course). Due to monthly fluctuations in EBP reach, we will calculate 12-month period pre/post EBP reach averages, which removes clustering of EBP reach observations *within* clinics over time in tests of specific aims.

Intraclass Correlation Coefficient (ICC). The ICC is the degree of similarity among EBP reach observations over time and within the same national VA health care system. We used the ICC to estimate the variance inflation factor/design effect for clustered data, a parameter for determining the number of clinics needed for tests of specific aims. Using R and the 'ICC' package,[166,167] we analyzed a 2-year database extracted from the CDW to estimate the ICC of EBP reach for clinics nested within the VA health care system, and repeated observations of clinics over time. Across EBP initiation and course, values were low,[168,169] ranging from low ICCs (< 0.0001) between clinics, to higher ICCs (~ 0.006) for repeated within-clinic observations. These ICC values indicate near independence,[168,169] and are consistent with pilot work identifying significant differences between clinics, and high within-clinic variability in EBP reach over time.[48]

Mediation analyses. Tests of Aim 2 are assessed on same between-clinic level: Clinics are assigned randomly to AF or PSD. We will use the clinic observation of team-means across arms as our mediator, and the clinic EBP reach as our outcome. Our mediator is adequately powered (d = .78).[135]

Average Clinic Cluster Size. ICC and average cluster size are key to determining CRT power, accounting for variance inflation due to clustering. Our pre/post EBP reach measure is the 12-month period average of 12 monthly EBP reach observations. In our 2-year CDW database, median clinic size was ~800 unique patients/month summed across diagnostic cohorts (depression + PTSD + OUD + AUD diagnoses).

Power analyses. We used 'CRTSize'[168,169] in R to calculate the number of clinics necessary to balance type 1 error (alpha), type 2 error (beta), and power, for a CRT with a binary outcome (difference in proportions). Omnibus EBPsy/EBPharm *initiation* (Aim 1a) is the limiting power analysis.[168,169] *Effect size*: We expect a 5% *initiation* increase in PSD clinics (from ~5 to 10%), which would meaningfully double omnibus EBP reach to exceed the national median. We expect little/no change in AF clinics (~ 5 to 5%). Thus, we expect PSD intervention will to lead to a 5% difference in EBP reach between PSD and AF arms. With 11 clinics/arm we have power = 0.80 to detect a 5% difference in EBP initiation (two-tailed test, alpha = 0.05, average clinic/cluster size of 800, and ICC = 0.02). However, we will use an even 12 clinics/arm, better for stratified block randomization.[155] Power analyses used ICC = .02, because our STS team-mean measures should become more similar after PSD/AF, leading to higher ICC values than in our CDW clinic population database.[57]

Attrition. Attrition of providers or patients will not impact analyses for specific aims. Provider PSD participation will be tracked as a PSD fidelity check, but patient attrition is included by definition in EBP reach measures, and use of provider attrition in team-average (mean) survey measures means the only loss of data would be due to loss of an entire clinic team. Our R21 pilot testing indicates it is unlikely that care teams will attrit. Should clinics attrit after randomization, they will be included in intent-to-treat analyses using CDW data.

IRB PLAN

Our local IRB is the Stanford University IRB and the Office of Research at VAPAHCS.

The study team at the National Center for PTSD will manage the R01 trial protocol from the campus of the VAPAHCS. All clinics enrolled in this R01 study will complete the same randomization protocol and participate in either audit and feedback (AF) or in participatory system dynamics (PSD). These interventions and all trial records and data will be managed centrally from one lead site at NCPTSD/VAPAHCS.

PROTECTION OF HUMAN SUBJECTS

This Human Subjects Research involves an NIH-Defined Phase III Clinical Trial.

1. Risks to Human Subjects

a. *Human Subjects Involvement, Characteristics, and Design*

Human Subjects Involvement and Characteristics. Electronic data from the VA Corporate Data Warehouse (CDW) will be used to evaluate specific aims and comprise the *first of two* human subjects components to our study. Due to our proposal using existing data during this project, risks associated with this study to individual patients who use mental health services are minimal. There will be no interaction with current patients for the purposes of research. Patients will not be asked to sign a consent form. No new data will be collected beyond data generated during routine care. All individual patient data will stay on servers behind the VA firewall to prevent any potential risk for loss of confidentiality of protected health information. Data inputs in the models will be de-identified team aggregates and will not be individually identifiable.

Staff/Stakeholder Involvement. This project is a collaboration between the Principal Investigator and Co-Investigator team, and the leadership, front-line providers, and staff in the VA outpatient service system. Comprising the *second of two* human subjects components of our study, Mental Health staff will be engaged in the Audit and Feedback (AF) 'Team Feedback' and Participatory System Dynamics (PSD) 'Modeling to Learn' team trainings (i.e., they will participate in their actual team/workgroup) for no-cost, with continuing education credits provided by VA Employee Education Services toward licenses in psychiatry, psychology, social work, nursing, certified peer support specialty, and counseling. These educational trainings are commonly provided in VA and are necessary for maintenance of state licensure and VA hospital privileges. During workshop sessions, frontline teams of providers will partner to evaluate improvement scenarios via simulation and identify quality improvement changes in their clinics/teams.

Over the course of the proposed study, this research project expects to involve eight regional health systems randomized to either AF or PSD, in which the Medical Director has identified at least three clinics for participation in training (4 regional health systems/12 clinics per arm; 24 clinics total). We expect approximately 30 staff to participate in each outpatient clinic. Across 24 clinics that will include 720 staff. AF/PSD workshops with staff will be held during regularly scheduled staff meetings or team huddles. We anticipate that participation in PSD will include approximately two workshop hours per month for six months with optional self-directed learning. Participation in AF activities will include two hours in the first month, and less than one self-directed hour per month during months two through six. Participation will occur during work hours and require clinic manager workload credit approval to staff for participation. Staff will receive 12 continuing education credit hours for PSD and two continuing education credit hours for AF. Since randomization occurs at the VA regional health system level, the training opportunity will not vary among co-workers. Local staff will either all receive AF or all receive PSD. The leaders and staff of the outpatient service system helped to shape the goals of this

study. Veteran staff with lived experience using the mental health and addiction service systems, who now work as VA patient navigators, will continue to participate and shape the development of the project through all phases, including their role as workshop co-facilitators. These certified peer specialists from the Veteran Advisory Partnership for Operations and Research complete CITI Training.

b. *Sources of Materials*

Sources of Materials – Patient Human Subjects. Administrative data and VA information systems will be used to evaluate and compare staffing allocations, patient referral flows, appointment timing and type, and pharmacy records. These electronic data will be drawn from the regional VA Corporate Data Warehouse (CDW).

Sources of Materials – Staff/stakeholder Human Subjects. We will collect anonymized, qualitative and quantitative data from mental health staff. Routine data about team care patterns, and information about staff decisions regarding change will be collected throughout the project as part of PSD and AF activities. Licensure accreditation bodies and VA employee education services each have educational course evaluation requirements that will not comprise research data, but will be required for staff to obtain licensure credits.

c. *Potential Risks*

The primary risk to human subjects/mental health patients is associated with potential breaches of confidentiality of patient health records. In addition, staff who participate may feel uncomfortable about the review of data and our focus on the performance of the mental health delivery system.

2. Adequacy of Protection Against Risks

a. Recruitment and Informed Consent

The research involves no more than minimal risk to the patient participants. Patient data will be extracted from VA administrative datasets. These data are collected during routine mental health care. We will seek a waiver of consent and a HIPAA waiver. We will not recruit, conduct informed consent, or interact with patients. Precautions will be taken to ensure that confidentiality is maintained. Individually identifiable patient data will not be synthesized in models. All files related to study data will be password protected and will only be accessible by those working directly on the study.

This research involves no more than minimal risk to the provider participants. AF is already standard practice in VA. PSD is an augmented version of AF for achieving greater improvements in quality. Eligible VA regional health systems will be below the overall VA national quality median. Regional VA health systems must have the Medical Director assent to randomization and have identified three clinics within their regional system that are willing to participate. All the clinics from the regional system, and the teams that comprise those clinics, will receive the same implementation strategy (Audit and Feedback or Participatory System Dynamics).

Recruitment will occur via the VA offices responsible for ensuring high-quality EBP delivery: Office of Mental Health and Suicide Prevention (OMHSP), National Center for PTSD (NCPTSD), Quality Enhancement Research Initiative (QUERI), Evidence-based Practice (EBP) Coordinators, and Regional Mental Health Leads. These national groups of directors, program leads and managers, each have listservs and regular meetings. We will use email and online/in-person presentations to introduce the AF/PSD team training opportunities, for the purposes of recruitment. The R01 study team includes co-investigators and advisors from OMHSP, NCPTSD, EBP Coordinators and Regional Mental Health Leads.

Providers will provide assent or implied consent to participate in the AF or PSD training workshop. Providers will be informed in online training materials that the team AF/PSD training opportunity is designed to improve the quality of addiction and mental health care with the opportunity to receive continuing education credits toward licensure. They will be informed the training is voluntary, and of the possible risks, such as discomfort talking about quality issues and team care coordination issues. Providers will also be informed of potential benefits, such as the ability to improve patient care and provider quality of work-life. Staff do not need to participate for continuing education credit to be involved in this team learning opportunity.

Staff online/written informed consent will be obtained for completion of a 52-item survey at baseline and an 83-item survey at six months for the purposes of research. Staff will be informed of the potential knowledge to be gained from the surveys. Staff will be informed of the opportunity to stop survey participation at any time without penalty. Any new findings developed during the course of the study, which may relate to staff willingness to continue to participate will be provided to staff. Providers will have the option not to participate in the survey measures or withdraw their participation at any time.

b. Protections Against Risk

Protection against Breaches of Confidentiality. Patient information in the VA administrative data systems will not be transported. Real-time aggregate data reports will be extracted from existing quality assurance algorithms leaving the data on VA servers. Individually identifiable patient data will not be synthesized in models. All files related to study data will be password protected and will only be accessible by those working directly on the study.

Protection against Staff Discomfort. Staff, by agreeing to participate in the voluntary team training, will be giving assent and permission to review health system data to improve EBP implementation and overall mental health and addiction service quality. This will be stated clearly at the beginning of project activities and participants may withdraw from the training or decline to discuss without penalty; withdrawing from the team trainings will in no way impact access to consultation or data services developed or made available as part of this study. Based on past research collaboration with outpatient stakeholders, we expect that staff will appreciate the opportunity to share their experiences, challenges, and concerns. However, should they experience distress as a result of participating in this research partnership, we will refer them to a member of our VA Office of Research and Development, and will notify the IRB. Co-Investigators and advisors on this grant have extensive experience working with and addressing the problems of VA staff (Drs. Lindley, Rosen, Kimerling, Collie, and Trafton).

3. Potential Benefits of the Proposed Research to Human Subjects and Others

The purpose of this project is to test the use of PSD against usual AF. Based on large bodies of AF and PSD research, we expect PSD to enable frontline mental health staff to better identify improvements to mental health delivery that increase the proportion of the patient population who receive high-quality, evidence-based care. We will learn about causes of limited EBP reach, the explanation for findings regarding AF/PSD effectiveness, and we will learn how generalizable the PSD models are across a wide range of clinics/regional health systems. This project should improve Veterans' health and well-being by improving the quality of their care. It is possible that PSD/AF activities will also improve providers' quality of work-life on their teams as improvements are identified. The electronic data systems synthesized in system dynamics models will also enable us to prepare to expand this benefit more widely beyond VA through future research through online public dissemination of models, code and training resources. Finally, this project is designed to maximally support and increase the capacities of stakeholders in outpatient services delivering care via mental health, and therefore we aim to build the skills and knowledge of staff who are participating.

4. Importance of the Knowledge to be Gained

We're proposing a Phase III Clinical Superiority design with falsifiable hypothesis tests. We're testing the most commonly used, but often ineffective strategy in the world (AF), against a very rigorous theory-based approach (PSD) with a 60-year track record of effectiveness for improving organizations/business. Confirmatory effectiveness (Aim 1), causality (Aim 2) and generalizability (Aim 3) hypotheses are relatively rare in the field of implementation science. The ultimate anticipated benefit of this project lies in its potential to identify consistent ways to increase the reach of evidence-based practices (EBP) to patients in need of services. Toward this end, study aims are designed to inform replication of the use of PSD should R01 findings warrant further study. Our project activities are the first step toward creating a paradigm for ongoing quality improvements in health service delivery. Due to our use of VA nationwide data extraction approaches, this method can be scaled and applied to other VA health specialties beyond mental health. Use of standard data definitions for coding patient diagnoses or patient-provider encounters (visits) makes the study ready for replication in any U.S. healthcare system. This program of research could improve the quality of care in the future for large populations of patients and help the VA and other health systems to make better use of existing resources (i.e., staffing) to provide highly effective treatments.

RECRUITMENT AND RETENTION PLAN

Selection of Study Clinics – Inclusion/Exclusion Criteria

We will randomize and complete the participatory system dynamics '*Modeling to Learn'* (PSD) or audit and feedback '*Team Feedback'* (AF) virtual workshop with VA divisions and community-based outpatient clinics (CBOCs) or 'clinics.' Eligible clinics will be from regional VA health systems below the overall VA quality median (as assessed by the Strategic Analytics for Improvement and Learning or SAIL). This includes 3 of 8 SAIL measures associated with four evidence-based psychotherapies and three evidence-based pharmacotherapies for depression, PTSD, and opioid use disorder in (see *Table 1* in *Strategy*).

Inclusion criteria balance sensitivity and specificity in identifying clinics from lower performing health care systems. We focus on the lower half of our national distribution in EBP reach, which is the relevant population for inference from the R01 study sample, and for interpreting R01 results within the literature.

Recruitment of Study Clinics

Once identified via SAIL, recruitment will occur via the networks and training programs of the VA Central Office quality programs run by the study team. These offices and programs include the Office of Mental Health and Suicide Prevention (OMHSP), the National Center for PTSD (NCPTSD), Dissemination and Training Division, the NCPTSD National Mentorship Program, the NCPTSD Practice-based Implementation Network, the OMHSP EBP Coordinator program, the OMHSP Technical Assistance Specialists, OMHSP Program Evaluation Resource Center, OMHSP Psychotropic Drug Safety Initiative, VA Academic Detailing Program, the Behavioral Health Integrated Plan (BHIP) Coaches Program, and VISN Mental Health Lead networks (Veterans Integrated Service Network [VISN] - multi-state regional health care system networks).

Clinics must have regional health care system director assent to randomization. Analyses will control for clinics nested within VA. Given known interest in participatory system dynamics among health care systems due to the activities of our quality improvement programs, recruiting from these networks should be highly feasible. Feasibility is further enhanced by 18 months total pre/post flexibility (see Section 2.7 of Human Subjects and Clinical Trials Information - *Timeline, Table 10*).

Engagement and Retention of Teams in Study Clinics

AF participants will engage during two regular team meetings in 1 month (2 hours), and receive weekly emails for 6 months; PSD participants will engage during two regular team meetings over 6 months (12 hours), with weekly emails for 6 months (Strategy *AF/PSD Workshop Training* and *AF Dynamic Data Tools*). AF/PSD will occur during normal meetings, substituting only the activities used to improve quality objectives.

Continuing Education Licensure Credit. All frontline addiction and mental health disciplines will have the opportunity to participate for licensure credit provided by VA Employee Education Services. This includes

psychiatrists, psychologists, social workers, nurses, counselors and certified peer support specialists. Providers randomized to the audit and feedback arm will have the opportunity to receive two hours of licensure credit for two facilitated team meetings. Providers randomized to participatory system dynamics will have the opportunity to receive twelve hours of licensure credit for twelve facilitated team meetings over six months.

Attrition. Attrition of providers or patients will not impact analyses for specific aims. Provider PSD participation will be tracked as a PSD fidelity check, but patient attrition is included by definition in EBP reach measures, and use of provider attrition in team-average (mean) survey measures means the only loss of data would be due to loss of an entire clinic team. Our R21 pilot testing indicates it is unlikely that care teams will attrit. Should clinics attrit after randomization, they will be included in intent-to-treat analyses using CDW data.

INCLUSION OF WOMEN, MINORITIES, AND CHILDREN

Inclusion of Women and Minorities:

The proposed research will be open to adult men and women of all ethnic and racial backgrounds. Women and members of minority groups from *two populations* will be included in the proposed R01 study: the VA patient population served in outpatient mental health and addiction services (existing VA health system data) and the VA outpatient mental health and addiction provider population (survey data; see second page). Our trial plans are inclusive for both patients and providers.

The R01 patient and provider samples will be representative of VA patient and provider populations. The VA patient population skews male. We address sex/gender, race, and ethnicity in our proposed trial design with inclusive eligibility criteria. All VA addiction and mental health patients with a primary diagnosis of depression, PTSD, alcohol use disorder and opioid use disorder are eligible. All VA addiction and mental health providers mapped to a care team in a study clinic are eligible. We will randomize provider participants to each arm.

Our evaluation of the trial outcome measure is drawn from existing health care records (see *Patients* below).

Prior studies neither support nor negate the potential for significant differences in participatory system dynamics or audit and feedback effectiveness for improving the reach of evidence-based practices among sex/gender, racial or ethnic subgroups. Our plans for valid analysis include reporting results by subgroup, and exploring for significant differences in our outcome of EBP reach (patients), and our mediator, systems thinking (providers). The targeted/planned distribution of subjects by sex/gender, racial, and ethnic groups for each proposed sample is provided in two Targeted/Planned Enrollment Tables below.

Patients. By VA policy, all VA patients should have equal access to EBPs. Existing health system data (e.g., means/median number of patients receiving specific services, means/median of scheduled clinic appointments) from a 2-year cohort of patient services data will be available for audit and feedback to frontline staff in order to achieve higher quality care (arm 1), or available for review by frontline staff and synthesized in participatory system dynamics models for simulation testing (arm 2). Analyses to test specific aims include a 24-month observation of health services delivery (12 months pre-/12 months post-) across 24 clinics.

Enrollment - Unique Patients in Multisite Mental Health Cohort					
	Not Hispanic or Latino		Hispanic or Latino		Total
Racial Categories	**Female**	**Male**	**Female**	**Male**	
American Indian/Alaska Native	314	3,157	50	504	4,025
Asian	64	649	10	104	827
Native Hawaiian or Other Pacific Islander	35	351	6	56	448
Black or African American	291	2,929	46	468	3,734
White	1,039	10,455	166	1,670	13,330
More than One Race					
Total	1,743	17,541	278	2,802	22,364

Providers. Teams (typically 4-6 staff) of frontline multidisciplinary mental health and addiction staff from 24 participating clinics will be randomized to six months of team audit and feedback or six months of team participatory system dynamics training, each designed for improvement in delivery of high-quality, evidence-based care. Pre-post provider survey measures will be collected at baseline and at the end of the improvement strategy (six months). All addiction and mental health providers mapped to a care team in our study clinics will have the opportunity to participate.

Provider Demographics					
	Not Hispanic or Latino		Hispanic or Latino		Total
Racial Categories	**Female**	**Male**	**Female**	**Male**	
American Indian/Alaska Native	8	5	1	1	15
Asian	32	20	3	2	57
Native Hawaiian or Other Pacific Islander	8	5	1	1	15
Black or African American	78	52	9	6	145
White	245	163	27	18	453
More than One Race	19	13	2	1	35
Total	390	258	43	29	720

Inclusion of Children:

All participants will be 18 years of age or older. There will be no children involved. This is a study of adult outpatient mental health and addiction services in the VA. The purpose of the study is to understand how to expand the reach of evidence-based psychotherapies and evidence-based pharmacotherapies determined to be effective for adult diagnosis with depression, PTSD, alcohol or opioid use disorder. The study will take place in the national VA health care system, which serves adult male and female patients.

This trial does not enroll patients and will not interact with patients for the purposes of research. Rather the evidence-based practice outcome measure will be observed in existing VA electronic health record systems.

The comparators in this two-arm cluster randomized trial are participatory system dynamics and audit and feedback. These two interventions are each designed to improve evidence-based practice implementation in the VA health care system by intervening with the professional (adult) frontline addiction staff of the VA.

Therefore, due to the trial target to expand the reach of adult treatments (i.e., evidence-based practices), and due to the focus on intervention with professional health care staff, participation of children in this trial is not scientifically or ethically justified.

STRUCTURE OF THE STUDY TEAM

Serving more than 8.9 million Veterans each year, VA is the largest health care system in the U.S., providing care at 1,053 outpatient clinics. Due to our participatory system dynamics (PSD) philosophy of science, interdisciplinary PSD application, and the scale of VA, a large team with a range of investigator, methodologist, and advisory expertise, ensures R01 success.

PI and Co-Investigators NCPTSD principal investigator and co-investigators (*Table 8*) have collectively facilitated many, multi-site implementation studies. R01 feasibility is further enhanced by our established partnerships. Eleven of seventeen partners collaborated effectively as R21 co-investigators or NCPTSD-funded PSD development partners.[48] Three new partners were added to our R01 team due to study relevant expertise with implementation science in addiction and mental health (McGovern), cluster randomized implementation studies (Wiltsey Stirman) and scaling system dynamics trainings (Snyder).

Advisory and Safety Monitoring Board In addition to our VAPOR partners, four study advisors lead national VA improvement initiatives: VA system organization (Rust), SAIL audit and feedback quality measures in mental health (Trafton), national evidence-based psychotherapy coordinator program (Collie) and the national evidence-based pharmacotherapy program (Wiechers). Rust and Trafton (R21 co-investigators) worked with us to develop *Modeling to Learn* at scale in VA. The study relevant expertise and contributions of each R01 study team member or partner is listed below in *Table 8.*

During the R21, we collaborated effectively across our nationally distributed team through regular meetings on the Lucid Meetings platform, project management via Basecamp 2.0, an open data science workflow on GitHub, and an integrated set of servers behind the VA firewall. We also benefit from co-location of NCPTSD, Office of Mental Health and Suicide Prevention (OMHSP), Veterans Advisory Partnership for Operations and Research (VAPOR) ,and Stanford partners at VAPAHCS (see included letters of support). NCPTSD, Georgia Health Policy Center (GHPC) and VAPOR staff will co-facilitate '*Modeling to Learn'* workshops. NCPTSD staff will provide post-workshop technical assistance.

Table 8. R01 Study Team & PSD Partners

Name	Role	Organization	Responsibilities during the R01 Trial	Specific Contributions
Zimmerman*+	PI	NCPTSD	Implementation, PSD, Mediation, Qualitative	Aims 1, 2 & 3
Lounsbury*	co-I	Einstein	Participatory System Dynamics (PSD)	Aims 1, 2 & 3
McGovern+	co-I	Stanford	Addiction, Mental Health, Implementation	Aims 1 & 2
Rosen*+	co-I	NCPTSD	EBP Dissemination, VA Multi-site Studies	Aims 1 & 2

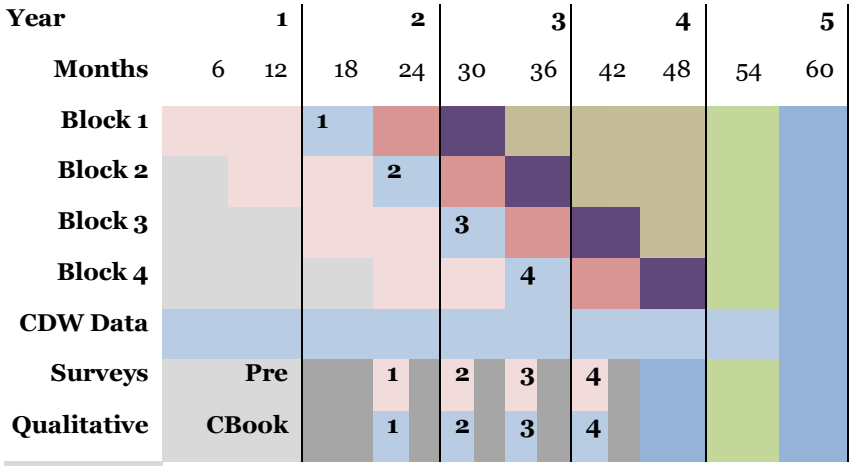
Kimerling*+	co-I	NCPTSD	Health Services Research, Staff Engagement	Aims 1 & 2
Wiltsey Stirman+	co-I	NCPTSD	Implementation Science/CRT	Aims 1 & 2
Lindley*+	co-I	VAPAHCS	Frontline Management	Field Expertise
Snyder	co-I	GHPC	PSD Facilitation Administration	PSD Trainings
Branscomb*	co-I	GHPC	PSD Facilitation/Facilitation Training	PSD Trainings
Hong*	key personnel	NCPTSD	CRT statistician, R and GitHub Expertise	Aims 1 & 2
Holbrook*	key personnel	VERC	Industrial Engineer/CDW SQL Programmer	CDW data/SQL code
Azevedo*+	key personnel	VAPAHCS	Qualitative Research and Coding	Aim 2 & Fidelity
Rust*	advisor	VERC	Systems Engineer/System Dynamicist	Aims 2 & 3
VAPOR*+	advisor	VAPAHCS	Peer Specialists: Patient Perspective	PSD Trainings
Trafton*+	advisor	OMHSP	SAIL Measures, VA Policy	AF/SAIL Code
Collie	advisor	OMHSP	VA VISN and EBP MH Leadership	AF/Clinic Engagement
Wiechers+	advisor	OMHSP	VA Psychotropic Drug Safety Initiative	AF/Clinic Engagement

Note: * Co-investigators on R21 and partners in NCPTSD development. + Co-located at VAPAHCS = VA Palo Alto Health Care System. NCPSTD = National Center for PTSD. Einstein = Albert Einstein College of Medicine. Stanford = Stanford University. OMHSP = Office of Mental Health and Suicide Prevention. VERC = Veterans Engineering Resource Center. GHPC = Georgia Health Policy Center. VAPOR = Veterans Advisory Partnership for Operations and Research. PSD = Participatory System Dynamics. AF = Audit and Feedback. CRT = Cluster Randomized Trial. CDW = Corporate Data Warehouse.

Timeline and Feasibility.

We propose 60 total months of study activities and 30 months of active PSD facilitation or post-training technical assistance (phases 3 and 4). This leaves 24 months of flexibility for delays across pre (phases 1 and 2) and post (phases 5 thru 8) activities (*Table 10*).

Table 10. Timeline of Proposed Study Activities



Phase 1	•IRB, Recruit, Randomize 24 Clinics: 12 per Arm & 6 per wave •Qualitative Codebook Developed/Adapted (Cbook)
Phase 2	•Pre 12-month EBP Reach; Team Data Websites & Models
Phase 3	•AF/PSD Trainings to Select Change Plans over 6 months
Phase 4	•AF/PSD Post-Training Technical Assistance
Phase 5	•***Aim 1*** AF/PSD Post 12-month period EBP reach •***Aim 2*** Qualitative Coding Analyses
Phase 6	•***Aim 2*** Survey Psychometric and Descriptive Analyses •*Explore* Online AF/PSD Sustained Use Monitoring
Phase 7	•***Aim 2*** Systems Thinking Mediation Analyses •***Aim 3*** Model Generality Analyses for AF Clinics
Phase 8	•Dissemination of Findings, Make Study Data Accessible •Disseminate PSD Workshop Scripts, Code & Models

We propose an implementation trial focused on the comparative effectiveness of two health care quality improvement strategies (participatory system dynamics vs audit and feedback). We will compare these two strategies to test their relative effectiveness for increasing the proportion of the outpatient addiction and mental health patient population that receives the highest quality, evidence-based standard of care.

Our two-arm trial focuses on changing provider care decisions to expand evidence-based treatments in routine care. Therefore, although our proposed R01 meets criteria for a phase III clinical trial, we expect that it is low risk with regard to patients and providers. We expect it to be low risk because we will not interact with patients for the purposes of research during this trial. Our focus is on the reach of evidence-based practices as measured in the VA electronic health record system. These psychotherapies and pharmacotherapies are referred to as "evidence-based practices" based on multiple randomized controlled trials, meta-analyses, and pragmatic effectiveness research studies.

This is a large national implementation science trial, but we are not examining the effectiveness of evidence-based practices. We are focused on expanding their reach among patients. There is no prior data to suggest that audit and feedback or participatory system dynamics approaches to quality improvement have significant adverse effects for patients or providers. Given this, we will take the following steps to ensure adequate data safety and monitoring.

Data Safety and Monitoring Plan

All investigators and project staff will complete necessary coursework regarding protection of human subjects and will receive certification from the Collaborative IRB Training Initiative (CITI). All investigators and project staff will remain current on VA privacy and information security trainings. We will also submit all procedures and documentation/definitions for electronic health record data collected to the relevant IRB (Stanford University) and VA Offices (VHA National Data Systems, VA Informatics and Computing Infrastructure, VA Office of Research and Development) for review and oversight.

We will maintain ongoing communication with our data safety and monitoring board and will regularly review data management procedures to identify and address unintended problems, and address any unlikely, but possible, adverse events.

Data Safety and Monitoring Board

Commensurate with the low risk of the trial, and commensurate with the size and complexity of our trial, our advisory board of VA quality improvement leaders in VA will comprise our data safety and monitoring board. One member runs the VA national audit and feedback program for the VA Office of Mental Health and Suicide Prevention (OMHSP) as the Director of the Program Evaluation Resource Center. A second member runs the national VA OMHSP quality improvement program for evidence-based pharmacotherapy as director of the Psychotropic Drug Safety Initiative. A third member runs the national VA quality improvement program for evidence-based psychotherapy as an OMHSP Technical Assistance Specialist. A fourth member runs national VA systems engineering quality improvement programs using EHR data. Finally, the Veterans Advisory Partnership for Operations and Research (VAPOR) provides ongoing input in study plans and evaluation from the Veteran patient perspective. This advisory board, will help to provide oversight and monitoring of the trial through the independent activities of their programs, and through regular communication with the study team.

We will meet regularly with our advisory board and will perform data safety and monitoring activities every six months. Monitoring activities will include review of study data in light of overall VA quality improvement data

with our complete multidisciplinary R01 study team. Our advisory board will provide assessments of trial progress based on the input of their independent teams of multi-disciplinary program evaluators, and will advise the R01 study team accordingly.

In addition to the advisory board, the R01 study team includes clinical trial experts, VA health services research experts, and clinicians who are experts in the evidence-based psychotherapies and pharmacotherapies that this trial seeks to make accessible to more VA patients.

DISSEMINATION PLAN

Registration of Clinical Trial

We will register this R01 trial and follow clinical trials reporting standards and policies via the ClinicalTrials.Gov website. NCPTSD, VA Palo Alto Health Care System (VAPAHCS), and Stanford University, all have internal policies in place to ensure that clinical trials registration and results reporting are in compliance federal clinical trials policy requirements.

Data, Facilitation Guides and Code

A central component of each arm of this participatory research project is to increase data transparency and accessibility among local frontline providers. Should participatory system dynamics prove to be superior and/or effective for improving reach of evidence-based practices, then dissemination of system dynamics modeling scripts, code, and models will be made available for use in other implementation contexts, and for replication by other implementation researchers.

Data. Datasets meeting VA standards for disclosure to the public will be made available within 1 year of publication. A de-identified, anonymized dataset will be created and shared. Final data sets underlying all publications resulting from the proposed research will be shared publicly.

Prior to distribution, a privacy officer will certify that all datasets contain no PII/PHI. Final data sets will be maintained locally, until VA and NIH enterprise-level resources become available for long-term storage and access. The VA Office of Research Development (ORD) will provide guidance on request and distribution processes. Those requesting data will be asked to sign a Letter of Agreement regarding use.

Publicly Available Participatory System Modeling Resources. In addition to these datasets, model SQL code for retrieving data from generic health record systems will be made publicly available online. The system dynamics model files and group modeling scripts used in the "*Modeling to Learn*" workshop series will be posted online for transparent, public use.

Publications

Publications from this research will be made available to the public through the National Library of Medicine PubMed Central website within one year after the date of publication, in accordance with guidance provided by NIH and VA ORD.

BIBLIOGRAPHY AND REFERENCES CITED

1. Elbogen EB, Wagner HR, Johnson SC, et al. Are Iraq and Afghanistan veterans using mental health services? New data from a national random-sample survey. *Psychiatr Serv*. 2013;64:134–141.
2. Hermes ED, Hoff R, & Rosenheck RA. Sources of the increasing number of Vietnam era veterans with a diagnosis of PTSD using VHA services. *Psychiatr Serv*. 2014;65:830–832.
3. Mott JM, Grubbs KM, Sansgiry S, et al. Psychotherapy utilization among rural and urban Veterans from 2007 to 2010. *J Rural Health*. 2015;31:235–243.
4. Hankin CS, Spiro III A, Miller DR, et al. Mental disorders and mental health treatment among US Department of Veterans Affairs outpatients: The Veterans Health Study. *Am J Psychiatry*. 2014.
5. Hoggatt KJ, Williams EC, Der-Martirosian C, et al. National prevalence and correlates of alcohol misuse in women Veterans. *J Subst Abuse Treat*. 2015;52:10–16.
6. Fulton JJ, Calhoun PS, Wagner HR, et al. The prevalence of posttraumatic stress disorder in Operation Enduring Freedom/Operation Iraqi Freedom (OEF/OIF) Veterans: A meta-analysis. *J Anxiety Disord*. 2015;31:98–107.
7. Karlin BE, & Cross G. From the laboratory to the therapy room: National dissemination and implementation of evidence-based psychotherapies in the U.S. Department of Veterans Affairs Health Care System. *Am Psychol*. 2014;69:19–33.
8. Karlin BE, Brown GK, Trockel M, et al. National dissemination of cognitive behavioral therapy for depression in the department of veterans affairs health care system: Therapist and patient-level outcomes. *J Consult Clin Psychol*. 2012;80:707–718.
9. Ruzek JI, Karlin BE, & Zeiss AM. Implementation of Evidence-Based Psychological Treatments in the Veterans Health Administration. In: McHugh RK, Barlow DH, eds. Dissemination of evidence-based psychological treatments. New York, NY: Oxford University Press. , 2012.
10. Eftekhari A, Ruzek JI, Crowley JJ, et al. Effectiveness of National Implementation of Prolonged Exposure Therapy in Veterans Affairs Care. *JAMA Psychiatry*. 2013;70:949.
11. Watts BV, Shiner B, Zubkoff L, et al. Implementation of evidence-based psychotherapies for posttraumatic stress disorder in VA specialty clinics. *Psychiatr Serv*. 2014;65:648–653.
12. Department of Defense, & Department of Veterans Affairs. The management of MDD Working Group. VA/DOD clinical practice guideline for management of major depressive disorder (MDD). , 2009.
13. Department of Veterans Affairs, & Department of Defense. VA/DoD Clinical practice guideline for the management of post-traumatic stress. , 2010.
14. Department of Veterans Affairs, & Department of Defense. VA/DoD Clinical practice guideline for the management of substance use disorders. , 2009.
15. Lemke S, Boden MT, Kearney LK, et al. Measurement-based management of mental health quality and access in VHA: SAIL mental health domain. *Psychol Serv*. 2017;14:1–12.
16. Harris AHS, Humphreys K, Bowe T, et al. Measuring the quality of substance use disorder treatment: Evaluating the validity of the Department of Veterans Affairs continuity of care performance measure. *J Subst Abuse Treat*. 2009;36:294–305.
17. Rosen CS, Matthieu MM, Wiltsey Stirman S, et al. A Review of Studies on the System-Wide Implementation of Evidence-Based Psychotherapies for Posttraumatic Stress Disorder in the Veterans Health Administration. *Adm Policy Ment Health Ment Health Serv Res*. 2016.
18. Seal KH, Maguen S, Cohen B, et al. VA mental health services utilization in Iraq and Afghanistan Veterans in the first year of receiving new mental health diagnoses. *J Trauma Stress*. 2010;:n/a-n/a.
19. Mott JM, Mondragon S, Hundt NE, et al. Characteristics of U.S. Veterans Who Begin and Complete Prolonged Exposure and Cognitive Processing Therapy for PTSD: Veterans in Evidence-Based Therapy for PTSD. *J Trauma Stress*. 2014;27:265–273.

20. Harpaz-Rotem I, & Rosenheck RA. Serving those who served: Retention of newly returning Veterans from Iraq and Afghanistan in mental health treatment. *Psychiatr Serv*. 2014.
21. Lin LA, Bohnert AS, Ilgen MA, et al. Outpatient provider contact prior to unintentional opioid overdose among VHA service users. *Psychiatr Serv*. 2015.
22. Harris AHS, Bowe T, Del Re AC, et al. Extended Release Naltrexone for Alcohol Use Disorders: Quasi-Experimental Effects on Mortality and Subsequent Detoxification Episodes. *Alcohol Clin Exp Res*. 2015;39:79–83.
23. Degenhardt L, Bucello C, Mathers B, et al. Mortality among regular or dependent users of heroin and other opioids: A systematic review and meta-analysis of cohort studies. *Addiction*. 2010;106:32–51.
24. Kaplan MS, Huguet N, McFarland BH, et al. Suicide among male veterans: a prospective population-based study. *J Epidemiol Community Health*. 2007;61:619–624.
25. Desai RA, Dausey DJ, & Rosenheck RA. Mental health service delivery and suicide risk: The role of individual patient and facility factors. *Am J Psychiatry*. 2014.
26. Gradus JL, Suvak MK, Wisco BE, et al. Treatment of posttraumatic stress disorder reduces suicidal ideation. *Depress Anxiety*. 2013;30:1046–1053.
27. Hysong SJ. Meta-analysis: Audit & feedback features impact effectiveness on care quality. *Med Care*. 2009;47:356.
28. Foy R, Eccles M, Jamtvedt G, et al. What do we know about how to do audit and feedback? Pitfalls in applying evidence from a systematic review. *BMC Health Serv Res*. 2005;5.
29. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: Effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev*. 2012;6.
30. Flottorp SA, Jamtvedt G, Gibis B, et al. Using audit and feedback to health professionals to improve the quality and safety of health care. *World Health Organ*. 2010.
31. Ivers NM, Sales A, Colquhoun H, et al. No more "business as usual"with audit and feedback interventions: Towards an agenda for a reinvigorated intervention. *Implement Sci*. 2014;9:14.
32. Colquhoun HL, Carroll K, Eva KW, et al. Advancing the literature on designing audit and feedback interventions: identifying theory-informed hypotheses. *Implement Sci*. 2017;12.
33. Colquhoun H, Michie S, Sales A, et al. Reporting and design elements of audit and feedback interventions: a secondary review. *BMJ Qual Saf*. 2017;26:54–60.
34. Landis-Lewis Z, Brehaut JC, Hochheiser H, et al. Computer-supported feedback message tailoring: Theory-informed adaptation of clinical audit and feedback for learning and behavior change. *Implement Sci*. 2015;10:12.
35. Atkinson J-A, Wells R, Page A, et al. Applications of system dynamics modelling to support health policy. *Public Health Res Pract*. 2015;25.
36. Diehl E, & Sterman JD. Effects of feedback complexity on dynamic decision making. *Organ Behav Hum Decis Process*. 1995;62:198–215.
37. Sterman JD. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Manag Sci*. 1989;35:321–339.
38. Cronin MA, Gonzalez C, & Sterman JD. Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. *Organ Behav Hum Decis Process*. 2009;108:116–130.
39. Sterman JD. Learning from evidence in a complex world. *Am J Public Health*. 2006;96:505–514.
40. Sterman JD. Learning in and about complex systems. *Syst Dyn Rev*. 1994;10:291–330.
41. Andersen DF, Vennix JA, Richardson GP, et al. Group model building: Problem structing, policy simulation and decision support. *J Oper Res Soc*. 2007;:691–694.
42. Vennix JAM. Group model building: facilitating team learning using system dynamics. Chichester ; New York: J. Wiley, 1996: 1-297.
43. Rouwette EAJA, Vennix JAM, & Mullekom T van. Group model building effectiveness: A review of assessment studies. *Syst Dyn Rev*. 2002;18:5–45.

44. Bendoly E. System dynamics understanding in projects: Information sharing, psychological safety, and performance effects. *Prod Oper Manag*. 2014;23:1352–1369.
45. Sterman JD. Does formal system dynamics training improve people's understanding of accumulation? *Syst Dyn Rev*. 2010;26:316–334.
46. Simon HA. Bounded rationality and organizational learning. *Organ Sci*. 1991;2:125–134.
47. Rahmandad H, Repenning N, & Sterman J. Effects of feedback delay on learning. *Syst Dyn Rev*. 2009;25:309–338.
48. Zimmerman L, Lounsbury DW, Rosen CS, et al. Participatory System Dynamics Modeling: Increasing Stakeholder Engagement and Precision to Improve Implementation Planning in Systems. *Adm Policy Ment Health Ment Health Serv Res*. 2016;43:834–849.
49. Oliva R. Model calibration as a testing strategy for system dynamics models. *Eur J Oper Res*. 2003;151:552–568.
50. Trafton JA, Greenberg G, Harris AHS, et al. VHA mental health information system: Applying health information technology to monitor and facilitate implementation of VHA Uniform Mental Health Services Handbook requirements. *Med Care*. 2013;51:S29-36.
51. Barlas Y. Formal aspects of model validity and validation in system dynamics. *Syst Dyn Rev*. 1996;12:183–210.
52. Hirsch G, Homer J, Trogdon J, et al. Using simulation to compare 4 categories of intervention for reducing cardiovascular disease risks. *Am J Public Health*. 2014;104:1187–1195.
53. Lich KH, Tian Y, Beadles CA, et al. Strategic Planning to Reduce the Burden of Stroke Among Veterans: Using Simulation Modeling to Inform Decision Making. *Stroke*. 2014;45:2078–2084.
54. Forrester, J.W. The model versus a modeling process. *Syst Dyn Rev*. 1985;:133–134.
55. Derickson R, Fishman J, Osatuke K, et al. Psychological safety and error reporting within Veterans Health Administration hospitals. *J Patient Saf*. 2015;11:60–66.
56. Edmondson A. Psychological safety and learning behavior in work teams. *Adm Sci Q*. 1999;44:350–383.
57. Campbell MJ, & Walters SJ. How to Design, Analyise and Report Cluster Randomised Trials in Medicine and Health Related Research. : Wiley, 2014: 1-264.
58. Forrester JW. System dynamics, systems thinking, and soft OR. *Syst Dyn Rev*. 1994;10:245–256.
59. Forrester JW. Some basic concepts in system dynamics. *Sloan Sch Manag Mass Inst Technol Camb MA*. 2009.
60. Hong G. Ratio of mediator probability weighting for estimating natural direct and indirect effects. In: Proceedings of the American Statistical Association, Biometrics Section. : American Statistical Association Alexandria, VA, 2010: 2401–2415.
61. Hong G, Deutsch J, & Hill HD. Ratio-of-Mediator-Probability Weighting for Causal Mediation Analysis in the Presence of Treatment-by-Mediator Interaction. *J Educ Behav Stat*. 2015;40:307–340.
62. Qin X, & Hong G. Causal mediation analysis in multi-site trials: An application of ratio-of-mediator-probability weighting to the Head Start Impact Study. *JSM Proc Soc Stat Sect*. 2014;:912–926.
63. Homer JB. Partial-model testing as a validation tool for system dynamics (1983). *Syst Dyn Rev*. 2012;28:281–294.
64. Rahmandad H, & Sterman JD. Reporting guidelines for simulation-based research in social sciences. *Syst Dyn Rev*. 2012;28:396–411.
65. Singer SJ, Moore SC, Meterko M, et al. Development of a short-form learning organization survey: The LOS-27. *Med Care Res Rev*. 2012;:1077558712448135.
66. Batorowicz B, & Shepherd TA. Measuring the quality of transdisciplinary teams. *J Interprof Care*. 2008;22:612–620.
67. Dolan ED, Mohr D, Lempa M, et al. Using a Single Item to Measure Burnout in Primary Care Staff: A Psychometric Evaluation. *J Gen Intern Med*. 2015;30:582–587.
68. Helfrich CD, Simonetti JA, Clinton WL, et al. The Association of Team-Specific Workload and Staffing with Odds of Burnout Among VA Primary Care Team Members. *J Gen Intern Med*. 2017.

69. Nelson KM, Helfrich C, Sun H, et al. Implementation of the Patient-Centered Medical Home in the Veterans Health Administration: Associations With Patient Satisfaction, Quality of Care, Staff Burnout, and Hospital and Emergency Department Use. *JAMA Intern Med*. 2014;174:1350.
70. Helfrich CD, Dolan ED, Simonetti J, et al. Elements of Team-Based Care in a Patient-Centered Medical Home Are Associated with Lower Burnout Among VA Primary Care Employees. *J Gen Intern Med*. 2014;29:659–666.
71. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. *Implement Sci*. 2017;12.
72. Sweeney LB, & Sterman JD. Bathtub dynamics: Initial results of a systems thinking inventory. *Syst Dyn Rev*. 2000;16:249–286.
73. Liberati EG, Ruggiero F, Galuppo L, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implement Sci.* 2017;12.
74. Centers for Medicare & Medicaid Services Alliance to Modernize Healthcare (CAMH). Independent Assessment of the Health Care Delivery Systems and Management Processes of the Department of Veterans Affairs (Volume 1: Integrated Report). .
75. Department of Veterans Affairs. Uniform mental health services in VA medical centers and clinics. Washington DC: Veterans Health Administration, 2008.
76. Powers MB, Halpern JM, Ferenschak MP, et al. A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clin Psychol Rev*. 2010;30:635–641.
77. Bradley R, Greene J, Russ E, et al. A Multidimensional Meta-Analysis of Psychotherapy for PTSD. *Am J Psychiatry*. 2005;162:214–227.
78. Powers MB, Zum V ouml rde Sive V ouml rding MB, & Emmelkamp PMG. Acceptance and Commitment Therapy: A meta-analytic review. *Psychother Psychosom*. 2009;78:73–80.
79. Walser RD, Karlin BE, Trockel M, et al. Training in and implementation of Acceptance and Commitment Therapy for depression in the Veterans Health Administration: Therapist and patient outcomes. *Behav Res Ther*. 2013;51:555–563.
80. Stewart MO, Raffa SD, Steele JL, et al. National dissemination of interpersonal psychotherapy for depression in veterans: Therapist and patient-level outcomes. *J Consult Clin Psychol*. 2014;82:1201–1206.
81. Cuijpers P, Geraedts AS, van Oppen P, et al. Interpersonal psychotherapy for depression: A meta-analysis. *Am J Psychiatry*. 2011.
82. Fournier JC, DeRubeis RJ, & Hollon SD. Antidepressant drug effects and depression severity: A patient-level meta-analysis. *JAMA*. 2010.
83. Knudsen HK, & Roman PM. The Diffusion of Acamprosate for the Treatment of Alcohol Use Disorder: Results From a National Longitudinal Study. *J Subst Abuse Treat*. 2016;62:62–67.
84. Knudsen HK, & Roman PM. Service delivery and pharmacotherapy for alcohol use disorder in the era of health reform: Data from a national sample of treatment organizations. *Subst Abuse*. 2016;37:230–237.
85. Pettinati HM, O'Brien CP, Rabinowitz AR, et al. The Status of Naltrexone in the Treatment of Alcohol Dependence: Specific Effects on Heavy Drinking. *J Clin Psychopharmacol*. 2006;26:610–625.
86. Anton RF, O'Malley SS, Ciraulo DA, et al. Combined pharmacotherapies and behavioral interventions for alcohol dependence: The COMBINE study: A randomized controlled trial. *JAMA*. 2006.
87. Barnett PG, Rodgers JH, & Bloch DA. A meta-analysis comparing buprenorphine to methadone for treatment of opiate dependence. *Addiction*. 2001.
88. West SL, O'Neal KK, & Graham CW. A meta-analysis comparing the effectiveness of buprenorphine and methadone. *J Subst Abuse*. 2001;12:405–414.
89. Hysong SJ. Theory and Evidence-Based Design of Audit and Feedback to Improve Quality of Care. 2017.
90. Hysong SJ, Best RG, & Pugh JA. Audit and feedback and clinical practice guideline adherence: Making feedback actionable. *Implement Sci*. 2006;1:5–3.

91. Hysong SJ, Teal CR, Khan MJ, et al. Improving quality of care through improved audit and feedback. *Implement Sci*. 2012;7:1–22.
92. Shiner B, D'Avolio LW, Nguyen TM, et al. Measuring Use of Evidence Based Psychotherapy for Posttraumatic Stress Disorder. *Adm Policy Ment Health Ment Health Serv Res*. 2013;40:311–318.
93. Glasgow RE, Vogt TM, & Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health*. 1999;89:1322–1327.
94. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health Ment Health Serv Res*. 2011;38:65–76.
95. Brownson RC, Colditz GA, & Proctor EK. Dissemination and Implementation Research in Health: Translating Science to Practice. Oxford, UK: Oxford, 2012.
96. Tabak RG, Khoong BS, Chambers DA, et al. Bridging research and practice: Models for dissemination and implementation research. *Am J Prev Med*. 2012;43:337–350.
97. Tabak RG, Khoong EC, Chambers D, et al. Models in dissemination and implementation research: useful tools in public health services and systems research. *Front Public Health Serv Syst Res*. 2013;2:8.
98. Damschroder LJ, Aron DC, Keith RE, et al. Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implement Sci*. 2009;4:50.
99. Powell BJ, Waltz TJ, Chinman MJ, et al. A refined compilation of implementation strategies: Results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci*. 2015;10:21.
100. Chambers D, R Glasgow, & K Strange. The dynamic sustainability framework: Addressing the paradox of sustainment amid ongoing change. *Implement Sci*. 2013;:117.
101. Breuer E, Lee L, De Silva M, et al. Using theory of change to design and evaluate public health interventions: a systematic review. *Implement Sci*. 2015;11.
102. Schierhout G, Hains J, Si D, et al. Evaluating the effectiveness of a multifaceted, multilevel continuous quality improvement program in primary health care: developing a realist theory of change. *Implement Sci*. 2013;8.
103. Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci*. 2015;10.
104. Sales A, Smith J, Curran G, et al. Models, Strategies, and Tools.: Theory in Implementing Evidence-Based Findings into Health Care Practice. *J Gen Intern Med*. 2006;21:S43–S49.
105. Damschroder LJ, & Lowery JC. Efficient Synthesis: Using Qualitative Comparative Analysis (QCA) and the CFIR across Diverse Studies. 2015.
106. Baumgartner M, & Thiem A. Often Trusted but Never (Properly) Tested: Evaluating Qualitative Comparative Analysis. *Sociol Methods Res*. 2017;:4912411770148.
107. Hug S. Qualitative Comparative Analysis: How Inductive Use and Measurement Error Lead to Problematic Inference. *Polit Anal*. 2013;21:252–265.
108. Krogslund C, Choi DD, & Poertner M. Fuzzy Sets on Shaky Ground: Parameter Sensitivity and Confirmation Bias in fsQCA. *Polit Anal*. 2015;23:21–41.
109. Braumoeller B. QCAfalsePositive: Tests for Type I Error in Qualitative Comparative Analysis (QCA). 2015.
110. Powell BJ, Proctor EK, & Glass JE. A Systematic Review of Strategies for Implementing Empirically Supported Mental Health Interventions. *Res Soc Work Pract*. 2014;24:192–212.
111. Williams NJ. Multilevel Mechanisms of Implementation Strategies in Mental Health: Integrating Theory, Research, and Practice. *Adm Policy Ment Health Ment Health Serv Res*. 2016;43:783–798.
112. Senge P. The Fifth Discipline. : Currency/Doubleday, 1990.
113. Sweeney LB, & Sterman JD. Bathtub dynamics: initial results of a systems thinking inventory. *Syst Dyn Rev*. 2000;16:249–286.
114. Hysong SJ, Knox MK, & Haidet P. Examining Clinical Performance Feedback in Patient-Aligned Care Teams. *J Gen Intern Med*. 2014;29:667–674.
115. Hysong SJ, Thomas CL, Spitzmüller C, et al. Linking clinician interaction and coordination to clinical performance in Patient-Aligned Care Teams. *Implement Sci*. 2015;11.

116. Richardson GP. Concept models in group model building: G. P. Richardson: Concept Models in Group Model Building. *Syst Dyn Rev*. 2013;29:42–55.
117. Richardson GP. Reflections on the foundations of system dynamics: Foundations of System Dynamics. *Syst Dyn Rev*. 2011;27:219–243.
118. Garcia HA, McGeary CA, Finley EP, et al. Burnout among psychiatrists in the Veterans Health Administration. *Burn Res*. 2015;2:108–114.
119. Garcia HA, Finley EP, Ketchum N, et al. A survey of perceived barriers and attitudes toward mental health care among OEF/OIF veterans at VA outpatient mental health clinics. *Mil Med*. 2014;179:273–278.
120. Garcia HA, Kelley LP, Rentz TO, et al. Pretreatment predictors of dropout from cognitive behavioral therapy for PTSD in Iraq and Afghanistan war veterans. *Psychol Serv*. 2011;8:1–11.
121. Senge, P.M. The fifth discipline: The art and practice of the learning organization. : Broadway Business, 2006.
122. Meadows DH. Thinking in systems: A primer. : Chelsea Green Publishing, 2012.
123. Senge PM, ed. The dance of change: the challenges of sustaining momentum in learning organizations, 1st ed. New York: Currency/Doubleday, 1999: 1-596.
124. Sterman JD. Business Dynamics: Systems Thinking and Modeling for a Complex World. : McGraw-Hill Education, 2000: 1-1008.
125. Rouwette EA, & Vennix JA. System dynamics and organizational interventions. *Syst Res Behav Sci*. 2006;23:451–466.
126. Akhnif E, Macq J, Idrissi Fakhreddine M., et al. Scoping literature review on the Learning Organisation concept as applied to the health system. *Health Res Policy Syst*. 2017;15.
127. Morecroft J, & Sherman J. Modeling for learning organizations. Portland OR: Productivity Press, 1994.
128. Argyris C, & Schön DA. Organizational learning: A theory of action perspective. *Reis*. 1997;:345–348.
129. Forrester JW. Industrial Dynamics. Cambridge, MA: MIT Press, 1961.
130. Simon HA. Models of Man: Social and Rational-Mathematical Essays on Rational Human Behavior in a Social Setting. Oxford, UK: Wiley, 1957.
131. Huz S, Andersen DF, Richardson GP, et al. A framework for evaluating systems thinking interventions: an experimental approach to mental health system change. *Syst Dyn Rev*. 1997;13:149–169.
132. Evans JSBT. Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annu Rev Psychol*. 2008;59:255–278.
133. Evans JSBT. In two minds: dual-process accounts of reasoning. *Trends Cogn Sci*. 2003;7:454–459.
134. Evans JSBT, & Stanovich KE. Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect Psychol Sci*. 2013;8:223–241.
135. Moore SM, Dolansky MA, Singh M, et al. The Systems Thinking Scale. 2010.
136. Hovmand PS. Community Based System Dynamics. New York, NY: Springer New York, 2014.
137. Minkler M, & Wallerstein N. Community-Based Participatory Research for Health: From Process to Outcomes. : John Wiley & Sons, 2011: 1-758.
138. Case AD, Byrd R, Claggett E, et al. Stakeholders' Perspectives on Community-Based Participatory Research to Enhance Mental Health Services. *Am J Community Psychol*. 2014;54:397–408.
139. Israel BA, Coombe CM, Cheezum RR, et al. Community-based participatory research: A capacity-building approach for policy advocacy aimed at eliminating health disparities. *Am J Public Health*. 2010;100:2094–2102.
140. Wallerstein NB. Using community-based participatory research to address health disparities. *Health Promot Pract*. 2006;7:312–323.
141. Aarons GA. Measuring Provider Attitudes Toward Evidence-Based Practice: Consideration of Organizational Context and Individual Differences. *Child Adolesc Psychiatr Clin N Am*. 2005;14:255–271.
142. Aarons GA, Hurlburt M, & Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Adm Policy Ment Health Ment Health Serv Res*. 2011;38:4–23.

143. Aarons GA, Ehrhart MG, Farahnak LR, et al. Aligning leadership across systems and organizations to develop a strategic climate for evidence-based practice implementation. *Annu Rev Public Health*. 2014;35:255–274.
144. Aarons GA, Green AE, Trott E, et al. The Roles of System and Organizational Leadership in System-Wide Evidence-Based Intervention Sustainment: A Mixed-Method Study. *Adm Policy Ment Health Ment Health Serv Res*. 2016.
145. Brimhall KC, Fenwick K, Farahnak LR, et al. Leadership, Organizational Climate, and Perceived Burden of Evidence-Based Practice in Mental Health Services. *Adm Policy Ment Health Ment Health Serv Res*. 2016;43:629–639.
146. Senge P, & Forrester JW. Tests for building confidence in system dynamics models. 1980;14:209–228.
147. Rahmandad H, & Sterman JD. Reporting guidelines for simulation-based research in social sciences: Reporting Guidelines for Simulation-Based Research. *Syst Dyn Rev*. 2012;28:396–411.
148. Hovmand PS, Andersen DF, Rouwette E, et al. Group model-building "scripts" as a collaborative planning tool: Scripts as a collaborative planning tool. *Syst Res Behav Sci*. 2012;29:179–193.
149. Andersen D, & Richardson G. Scripts for group model building. *Syst Dyn Rev*. 1997;13:107–129.
150. Hovmand, Peter, S., et al. Scriptapedia 4.0.6. 2013.
151. McCarthy JT, Hocum CL, Albright RC, et al. Biomedical System Dynamics to Improve Anemia Control With Darbepoetin Alfa in Long-Term Hemodialysis Patients. *Mayo Clin Proc*. 2014;89:87–94.
152. Gallaher E, Steensma DP, Chrisope TR, et al. Individualized Medicine and Biophysical System Dynamics: An Example from Clinical Practice in End Stage Renal Disease. 2011.
153. Homer J. Levels of evidence in system dynamics modeling. *Syst Dyn Rev*. 2014;30:75–80.
154. Oliva R, & Sterman JD. Death spirals and virtuous cycles. In: Handbook of Service Science. : Springer, 2010: 321–358.
155. Snow G. Package "blockrand." 2015.
156. Holowka DW, Marx BP, Gates MA, et al. PTSD diagnostic validity in Veterans Affairs electronic records of Iraq and Afghanistan Veterans. *J Consult Clin Psychol*. 2014;82:569–579.
157. Emmons KM, Weiner B, Fernandez ME, et al. Systems Antecedents for Dissemination and Implementation: A Review and Analysis of Measures. *Health Educ Behav*. 2012;39:87–105.
158. Rabin BA, Lewis CC, Norton WE, et al. Measurement resources for dissemination and implementation research in health. *Implement Sci*. 2015;11.
159. Lewis CC, Fischer S, Weiner BJ, et al. Outcomes for implementation science: An enhanced systematic review of instruments using evidence-based rating criteria. *Implement Sci*. 2015;10.
160. Glasgow RE. What does it mean to be pragmatic? Pragmatic methods, measures, and models to facilitate research translation. *Health Educ Behav*. 2013;40:257–265.
161. Glasgow RE, & Riley WT. Pragmatic measures: What they are and why we need them. *Am J Prev Med*. 2013;45:237–243.
162. Singh MK, Ogrinc G, Cox KR, et al. The Quality Improvement Knowledge Application Tool Revised (QIKAT-R): *Acad Med*. 2014;89:1386–1391.
163. Rosseel Y, Oberski D, Byrnes J, et al. Package "lavaan" Version 0.5-23.1097. 2017.
164. Shavelson RJ, & Webb NM. Generalizability theory: A primer. Newbury Park, CA: Sage., 1991.
165. Millsap RE, & Kwok O-M. Evaluating the Impact of Partial Factorial Invariance on Selection in Two Populations. *Psychol Methods*. 2004;9:93–115.
166. Wolak M. Package "ICC" Version 2.3.0. 2015.
167. Bates D, Maechler M, Bolker B, et al. Package "lme4" Version 1.1-13. 2017.
168. Rotondi MA. Package "CRTSize" Version 1.0. 2015.
169. Rotondi MA, & Donner A. Sample Size Estimation in Cluster Randomized Educational Trials: An Empirical Bayes Approach. *J Educ Behav Stat*. 2009;34:229–237.
170. Ventana Systems Inc. Vensim@ Version 6.3. 2014.

171. Hyndman, R.J., & Khandakar, Y. Automatic time series forecasting: The forecast package for R. *J Stat Softw*. 2008;26.
172. Hyndman R, O'Hara-Wild Mi, Bergmeir C, et al. Package "forecast" Version 8.0. 2017.
173. Mott JM, Hundt NE, Sansgiry S, et al. Changes in psychotherapy utilization among veterans with depression, anxiety, and PTSD. *Psychiatr Serv*. 2014.
174. Oliva EM, Trafton JA, Harris AHS, et al. Trends in Opioid Agonist Therapy in the Veterans Health Administration: Is Supply Keeping up with Demand? *Am J Drug Alcohol Abuse*. 2013;39:103–107.
175. Diaz M. Pasteur and parachutes: When statistical process control is better than a randomized controlled trial. *Qual Saf Health Care*. 2005;14:140–143.
176. Duclos A, & Voirin N. The p-control chart: a tool for care improvement. *Int J Qual Health Care*. 2010;22:402–407.
177. Ogrinc GS, & Headrick L. Fundamentals of health care improvement: A guide to improving your patients' care. : Joint Commission Resources, 2008.
178. Ogrinc G, Mooney SE, Estrada C, et al. The SQUIRE (Standards for QUality Improvement Reporting Excellence) guidelines for quality improvement reporting: explanation and elaboration. *Qual Saf Health Care*. 2008;17:i13–i32.
179. Buntin MB, Burke MF, Hoaglin MC, et al. The benefits of health information technology: A review of the recent literature shows predominantly positive results. *Health Aff (Millwood)*. 2011;30:464–471.
180. Raffa SD, Maciejewski ML, Zimmerman LE, et al. A System-Level Approach to Overweight and Obesity in the Veterans Health Administration. *J Gen Intern Med*. 2017;32:79–82. PMCID: PMC5359151
181. Kilbourne AM, Elwy AR, Sales AE, et al. Accelerating Research Impact in a Learning Health Care System: VA's Quality Enhancement Research Initiative in the Choice Act Era. *Med Care*. 2016.
182. Mabry PL, Olster DH, Morgan GD, et al. Interdisciplinarity and systems science to improve population health. *Am J Prev Med*. 2008;35:S211–S224.
183. Mabry PL, Milstein B, Abraido-Lanza AF, et al. Opening a window on systems science research in health promotion and public health. *Health Educ Behav*. 2013;40:5S–8S.
184. Luke DA, & Stamatakis KA. Systems science methods in public health: Dynamics, networks, and agents. *Annu Rev Public Health*. 2012;33:357–376.
185. Moore GF, & Evans RE. What theory, for whom and in which context? Reflections on the application of theory in the development and evaluation of complex population health interventions. *SSM - Popul Health*. 2017;3:132–135.
186. Monks T. Operational research as implementation science: definitions, challenges and research priorities. *Implement Sci*. 2015;11.
187. May CR, Johnson M, & Finch T. Implementation, context and complexity. *Implement Sci*. 2016;11.
188. Brainard J, & Hunter PR. Do complexity-informed health interventions work? A scoping review. *Implement Sci*. 2015;11.

APPENDIX

Measures

1. **Acceptability of Intervention Measure (AIM)**
2. **Intervention Appropriateness Measure (IAM)**
3. **Feasibility of Intervention Measure (FIM)**
4. **Patient Aligned Care Team Burnout Measure (PACT)**
5. **Learning Organization Survey (LOS-27)**
6. **Team Decision Making Questionnaire (TDMQ)**
7. **Systems Thinking Scale (STS)**
8. **Demographic Measures**
9. **Anonymous Feedback**

Acceptability of Intervention Measure (AIM)

Please answer these questions on a scale of 1 to 5, where 1 = completely disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = completely agree

1. [Participatory System Dynamics OR Audit and Feedback] meets my approval
2. [Participatory System Dynamics OR Audit and Feedback] is appealing to me.
3. I like [Participatory System Dynamics OR Audit and Feedback].
4. I welcome [Participatory System Dynamics OR Audit and Feedback].

Intervention Appropriateness Measure (IAM)

Please answer these questions on a scale of 1 to 5, where 1 = completely disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = completely agree

1. [Participatory System Dynamics OR Audit and Feedback] seems fitting.
2. [Participatory System Dynamics OR Audit and Feedback] seems suitable.
3. [Participatory System Dynamics OR Audit and Feedback] seems applicable.
4. [Participatory System Dynamics OR Audit and Feedback] seems like a good match.

Feasibility of Intervention Measure (IAM)

Please answer these questions on a scale of 1 to 5, where 1 = completely disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = completely agree

1. [Participatory System Dynamics OR Audit and Feedback] seems implementable.
2. [Participatory System Dynamics OR Audit and Feedback] seems possible.
3. [Participatory System Dynamics OR Audit and Feedback] seems doable.
4. [Participatory System Dynamics OR Audit and Feedback] seems easy to use.

Patient Aligned Care Team Burnout Measure

1. How many years of experience have you had working with your team? [Enter # of years]

2. Are you currently on more than one team (Yes/No)?

3. Has your team had any changes in, loss of, staff in the past 12 months (Yes/No)?

Please answer question 4 on a scale of 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Very Often, 5 = Always

4. How often does your team work extended hours?

Please answer question 5 on a scale of 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Very Often, 5 = Always

5. Overall, based on your definition of burnout, how would you rate your level of burnout at work?
    - I enjoy my work. I have no symptoms of burnout.
    - Occasionally I am under stress, and I don't always have as much energy as I once did, but I don't feel burned out.
    - I am definitely burning out and have one or more symptoms of burnout, such as physical and emotional exhaustion.
    - The symptoms of burnout that I'm experiencing won't go away. I think about frustration at work a lot.
    - I feel completely burned out and often wonder if I can go on. I am at the point where I may need some changes or may need to seek some sort of help.

Learning Organization Survey

Please respond to each item in terms of how descriptive it is of your clinic.

Never - - - Always
0 1 2 3 4

1. In this clinic, people value new ideas.
2. Differences in opinions are welcomed in this clinic.
3. In this clinic, people are open to alternative ways of getting work done.
4. People in this clinic are eager to share information about what doesn't work as well as to share information about what does work.
5. This clinic engages in productive conflict and debate during discussions.
6. In this clinic, we frequently identify and discuss underlying assumptions that might affect key decisions.
7. If you make a mistake in this clinic, it is often held against you.
8. This clinic experiments frequently with new product/service offerings.
9. This clinic experiments frequently with new ways of working.
10. This clinic frequently employs pilot projects or simulations when trying out new ideas.
11. This clinic has a formal process for conducting and evaluating experiments or new ideas.
12. Experienced employees in this clinic receive training when shifting to a new position.
13. Experienced employees in this clinic receive training when new initiatives are launched.
14. Newly hired employees in this clinic receive adequate training.
15. This clinic has forums for meeting with and learning from: Experts from outside the organization.
16. This clinic has forums for meeting with and learning from: Experts from other departments/teams/divisions.
17. This clinic has forums for meeting with and learning from: Customers/clients
18. This clinic regularly conducts post-audits, after-action reviews, and debriefings.
19. There is simply no time for reflection in this clinic.
20. In this clinic people are too busy to invest time in improvement.
21. This clinic frequently compares its performance to: Best-in-class organizations.
22. This clinic frequently compares its performance to: Other similar clinics.
23. This clinic consistently collects information on technological trends.

Please respond to each item in terms of how descriptive it is of your clinic.

Highly Inaccurate - - - - - - Highly Accurate
0 1 2 3 4 5 6 7

24. Clinic manager(s) establish forums for and provide time and resources for identifying problems and organizational challenges.
25. Clinic manager(s) establish forums for and provide time and resources for reflecting and improving on past performance.
26. Clinic manager(s) listen attentively.
27. Clinic mangers invite input from others in discussions.

Team Decision Making Questionnaire

When completing this questionnaire please consider your overall experience with Team Decision Making (TDM). Please read each question and answer according to the scale below.

7 = To a vast extent
6 = To a very great extent
5 = To a great extent
4 = To a moderate extent
3 = To a small extent
2 = To a very small extent
1 = Not at all
N/A = Not Applicable

To what extent does the [Participatory System Dynamics OR Audit and Feedback] process help you to . . .

1. obtain support in clinical decision making?
2. make consistent recommendations for all clients?
3. apply standards consistently across your team?
4. takes personal onus off decisions regarding prescriptions?
5. validate your clinical decisions?
6. apply policies consistently within your own caseload?
7. apply policies accurately?
8. provide support with colleagues' clinical decision making?
9. share innovative ideas?
10. obtain clinical advice?
11. become more competent?
12. share success?
13. keep current with knowledge regarding changing policies?
14. learn about application of new technology/strategies?
15. obtain various clinical perspectives?
16. keep current with equipment and new technology in this field of clinical practice?
17. develop effective problem solving?
18. ensure quality of services?
19. generate new ideas with colleagues?

Systems Thinking Scale

Please read each of the statements and answer with a number from the 5-point scale below that indicates frequency of agreement with the statement:

1 = Never
2 = Seldom
3 = Some of the time
4 = Often
5 = Most of the time

When I want to make an improvement. . .

1. I seek everyone's view of the situation.
2. I look beyond a specific event to determine the cause of the problem.
3. I think understanding how the chain of events occur is crucial.
4. I include people in my work unit to find a solution.
5. I think recurring patterns are more important than any one specific event.
6. I think of the problem at hand as a series of connected issues.
7. I consider the cause and effect that is occurring in a situation.
8. I consider the relationships among coworkers in the work unit.
9. I think that systems are constantly changing.
10. I propose solutions that affect the work environment, not specific individuals.
11. I keep in mind that proposed changes can affect the whole system.
12. I think more than one or two people are needed to have success.
13. I keep the mission and purpose of the organization in mind.
14. I think small changes can produce important results.
15. I consider how multiple changes affect each other.
16. I think about how different employees might be affected by the improvement.
17. I try strategies that do not rely on people's memory.
18. I recognize system problems are influenced by past events.
19. I consider the past history and culture of the work unit.
20. I consider that the same action can have different effects over time, depending on the state of the system.

Demographic Measures

1. **Are you Hispanic, Latino, Latina, Latinx?**

   [ ] Yes
   [ ] No
   [ ] Prefer not to say

2. **What is your racial background? (Select all that apply.)**

   [ ] American Indian / Alaska Native
   [ ] Asian
   [ ] Native Hawaiian or Other Pacific Islander
   [ ] Black or African American
   [ ] White
   [ ] Prefer not to say

3. **Which racial background do you identify with the most?**

   [ ] American Indian / Alaska Native
   [ ] Asian
   [ ] Native Hawaiian or Other Pacific Islander
   [ ] Black or African American
   [ ] White
   [ ] More than One Race
   [ ] Prefer not to say

4. **What is your gender? (Select all that apply).**

   [ ] Man
   [ ] Woman
   [ ] Non-binary
   [ ] Prefer not to say

Anonymous Feedback

1. **Please share any anonymous feedback that you have about *Modeling to Learn:***