



Studying the implementation of Zero Suicide in a large health system: Challenges, adaptations, and lessons learned

Edwin D. Boudreaux^{a,*}, Celine Larkin^b, Ana Vallejo Sefair^a, Eric Mick^c, Karen Clements^c, Lori Pelletier^d, Chengwu Yang^c, Catarina Kiefe^c

^a Departments of Emergency Medicine, Psychiatry, and Population and Quantitative Health Sciences, UMass Chan Medical School, Worcester, MA, USA

^b Departments of Emergency Medicine and Psychiatry, UMass Chan Medical School, Worcester, MA, USA

^c Department of Population and Quantitative Health Sciences, UMass Chan Medical School, Worcester, MA, USA

^d Connecticut Children's Hospital, Hartford, CT, USA

ARTICLE INFO

Keywords:

Suicide
Suicide prevention
Mental health
Quality improvement
Implementation science

ABSTRACT

Background: Suicide remains the 10th leading cause of death in the United States. Many patients presenting to healthcare settings with suicide risk are not identified and their risk mitigated during routine care. Our aim is to describe the planned methodology for studying the implementation of the Zero Suicide framework, a systems-based model designed to improve suicide risk detection and treatment, within a large healthcare system.

Methods: We planned to use a stepped wedge design to roll-out the Zero Suicide framework over 4 years with a total of 39 clinical units, spanning emergency department, inpatient, and outpatient settings, involving ~310,000 patients. We used Lean, a widely adopted a continuous quality improvement (CQI) model, to implement improvements using a centralize “hub” working with smaller “spoke” teams comprising CQI personnel, unit managers, and frontline staff.

Results: Over the course of the study, five major disruptions impacted our research methods, including a change in The Joint Commission's safety standards for suicide risk mitigation yielding massive system-wide changes and the COVID-19 pandemic. What had been an ambitious program at onset became increasingly challenging because of the disruptions, requiring significant adaptations to our implementation approach and our study methods.

Conclusions: Real-life obstacles interfered markedly with our plans. While we were ultimately successful in implementing Zero Suicide, these obstacles led to adaptations to our approach and timeline and required substantial changes in our study methodology. Future studies of quality improvement efforts that cut across multiple units and settings within a given health system should avoid using a stepped-wedge design with randomization at the unit level if there is the potential for sentinel, system-wide events.

1. Background

Despite increased public attention, suicide remains the 10th leading cause of death in the United States (US) [1]. In 2020, there were 44,834 deaths by suicide, about 123 per day [2]. Emergency department (ED) visits related to suicidal ideation or suicide attempts steadily increased for nearly two decades, approaching 1.1% of all visits in 2017 [3]. In March 2020, related to the COVID-19 pandemic, there was a brief

decrease in suicide-related ED visits; however, after April 2020, they began rising again [4].

Healthcare settings hold considerable promise for suicide prevention. However, many patients with suicide risk are not identified because standardized screening with validated measures is not routinely used by clinicians [5–9]. Even when suicide risk is identified, often care is inconsistent with best practice guidelines [10,11], and patients easily “fall through the cracks” during their transition across care settings.

Abbreviations: (ED), Emergency Department; (NIMH), National Institute of Mental Health; (CQI), continuous quality improvement; (SOS), System of Safety; (EHR), electronic health record; (GIS), Geographic Information Systems.

* Corresponding author. University of Massachusetts Medical School, Emergency Medicine, LA-189, 55 Lake Avenue North, Worcester, MA, 01655, USA.

E-mail addresses: Edwin.Boudreaux@umassmed.edu (E.D. Boudreaux), Celine.Larkin@umassmed.edu (C. Larkin), Ana.VallejoSefair3@umassmed.edu (A.V. Sefair), Eric.Mick@umassmed.edu (E. Mick), Karen.Clements@umassmed.edu (K. Clements), lpelletier01@connecticutchildrens.org (L. Pelletier), Chengwu.Yang@umassmed.edu (C. Yang), Catarina.Kiefe@umassmed.edu (C. Kiefe).

<https://doi.org/10.1016/j.conctc.2022.100999>

Received 14 March 2022; Received in revised form 12 August 2022; Accepted 8 September 2022

Available online 17 September 2022

2451-8654/© 2022 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Therefore, the Action Alliance for Suicide Prevention [12] and the National Institute of Mental Health (NIMH) [13] have promoted studying systems-based models such as Zero Suicide [14]. The Zero Suicide framework, summarized in Fig. 1, emphasizes seven Essential Elements of Suicide Care [14] to guide healthcare system transformation: (1) Lead, (2) Train, (3) Identify, (4) Engage, (5) Treat, (6) Transition, and (7) Improve. Zero Suicide has become the dominant systems-based suicide prevention model in the US, but we need to better understand its impact on outcomes and the barriers and facilitators for implementation. System-based models, such as Zero Suicide, have shifted from placing the responsibility to identify and mitigate suicide risk entirely on individual clinicians to emphasizing continuous quality improvement (CQI) approaches to support risk identification, assessment, intervention, and care transitions, and fostering a leadership-driven safety culture. Early evidence from the Henry Ford Health System [15,16] and others [17,18] suggests such systems-based models may reduce suicide by up to 75%, and larger implementation efforts continue to be reported [19,20].

Our NIMH-funded System of Safety (SOS) studies Zero Suicide implementation within the largest healthcare system in central Massachusetts representing a catchment area of over one million people. SOS was implemented in ED, inpatient, and outpatient settings; engaged medical and behavioral health units and clinicians; targeted adults and children; and supported integration and collaboration across the entire system to provide a 360-degree safety net for patients at risk for suicide. Here, we present an overview of the planned research methods, describe disruptions requiring adjustments in the original plan, and summarize the resulting revised methods. This non-linear story where “the best laid plans go awry” is emblematic of challenges to multi-component implementation efforts in busy, complex health systems, and the lessons learned on pivoting and adaptation can inform similar implementation research endeavors.

2. SOS: planned study design and procedures

Overview of the intervention. The Zero Suicide framework guided SOS (Fig. 1). Of its seven Elements, three – Lead, Train, Improve – are at the system level, intended to foster a safety culture, train clinicians, and use data to drive improvement; these elements aligned strongly with our CQI model, Lean (below). The other four Elements – Identify, Engage,

Treat, Transition – act at the patient level. We intended to use a hub-and-spoke team model (Fig. 2). A central System Hub of subject matter experts, Lean CQI engineers, and health system leaders would oversee the effort across all Waves, hospitals, and settings, while smaller Lean teams, or Spokes, of unit managers and frontline staff were to inform the adaptations and oversee implementation at the unit level.

Of the numerous Zero Suicide clinical processes [21], we planned to evaluate a subset: (1) standardized, evidence-based suicide risk screening, (2) personalized safety planning, (3) means restriction counseling, and (4) post-acute care transition facilitation. We chose these because they are all relevant to different levels of care (ED, inpatient, and outpatient) and subpopulations (youth, adults, medical and mental health patients); are consistent with best practice suicide prevention [22]; align well with The Joint Commission’s standards [23]; were used in previous pragmatic clinical trials (i.e., ED-SAFE 1 and 2) [24,25]; and can be mapped to discrete data elements in the health system’s electronic health record (EHR), hence making them measurable with such data.

Just training clinicians to adhere to protocols without the larger

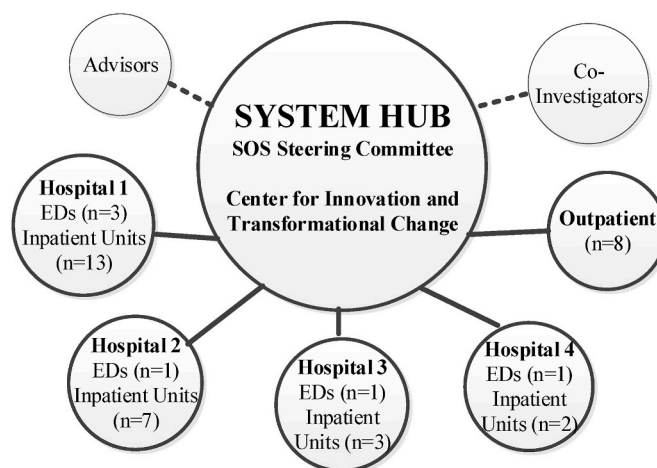


Fig. 2. Hub and spoke Team Design.

ZERO SUICIDE 7 ESSENTIAL CARE ELEMENTS	IDENTIFICATION OF EMPIRICALLY SUPPORTED OR BEST PRACTICE GUIDELINES	IMPLEMENTATION AND MAINTENANCE THROUGH LEAN CQI	EVALUATION AND OUTCOMES
<p>1. Lead - Create a leadership-driven, safety-oriented culture committed to reducing suicide.</p> <p>2. Train - Develop a competent, confident, and caring workforce.</p> <p>3. Identify - Systematically identify and assess suicide risk among people receiving care.</p> <p>4. Engage - Ensure every individual has a timely and adequate pathway to care</p> <p>5. Treat - Use effective, evidence-based treatments directly targeting suicidal thoughts and behaviors.</p> <p>6. Transition - Provide continuous contact and support, especially after acute care</p> <p>7. Improve - Apply data-driven quality improvement to inform system changes leading to improved patient outcomes and better care for those at risk.</p>	<p>1. Identify: Work to identify empirically supported strategies to accomplish each of the essential elements of care relevant to direct patient care (3 – 6). Where direct empirically supported strategies are lacking, identify best practice guidelines.</p> <p>2. Setting: Repeat process for each level of care (ED, inpatient, primary care) and discipline (medical, behavioral health).</p> <p>3. Translate into guidance: SOS study team works directly with the Hub and Spokes to build for each setting: (a) required performance elements for practice protocols and (b) performance goals and metrics.</p>	<p>1. Building and tailoring protocols. System Hub works with Local Hubs and Spokes representing care locations.</p> <p>2. Testing and implementation. Data generated by System Hub is used to improve clinician adherence, promote site accountability, measure implementation progress, and drive protocol adjustments.</p> <p>3. Collaborative learning. Quarterly and as needed collaborative learning sessions occur to promote cross-hospital, cross-setting knowledge transmission and collaboration.</p> <p>4. Maintenance. Active maintenance and continued support to ensure that implementation success is translated into enduring change.</p>	<p>1. Electronic health record (EHR) data. Data elements matching the 6 performance elements and patient outcomes are identified and tracked across entire study. These data are primary source for statistical analysis.</p> <p>2. Structured chart abstractions. Because EHR data even from structured fields can miss important details, structured chart abstractions including EHR free text by trained research staff is performed on randomly selected charts.</p> <p>3. Vital statistics. Work with the MA Department of Public Health to identify all probable and definite suicides in the greater Worcester region to perform failure mode analyses.</p> <p>4. Data is regularly summarized and used by the Hubs and Spokes to monitor and improve performance.</p>

Fig. 1. Relation between zero suicide, best practice, lean CQI, and evaluation and outcomes.

context of systems-change and support, colloquially dubbed the “train and pray” approach, is usually insufficient [26,27]. For SOS, we used Lean [28], a widely adopted CQI model, with strategies as in Fig. 3 [29, 30]. Lean integrates leadership engagement, continuous process changes, culture change, and accessible tools. It emphasizes measurement and reporting objective results.

SOS represents an unparalleled opportunity to study potential moderators, mediators, and mechanisms of action across an entire healthcare system including: (a) frontline clinician knowledge, attitudes, self-efficacy, and practice related to suicide care; (b) health system and departmental leadership support for these efforts; (c) enabling attributes, such as embedded mental health clinicians; and (d) outpatient behavioral health treatment after an acute care episode. These mechanisms, important in our team’s previous efforts, enjoy broad support in the CQI literature [31–33], and in implementation science models, e.g. the Practical, Robust Implementation, and Sustainability Model (PRISM) [34] and the Consolidated Framework for Implementation Research (CFIR) [35].

Our hypotheses were as follows:

- **Hypothesis 1:** Likelihood of suicide risk screening at the time when a patient enters the study will increase monotonically with time since study initiation.
- **Hypothesis 2:** Likelihood of suicide risk identification (patient outcome) at the time when a patient enters the study will increase monotonically with time since study initiation.
- **Hypothesis 3:** As SOS is implemented across more settings and clinical units, the likelihood of receiving a best practice suicide-prevention intervention by a clinician will increase.

Study phases. A stepped wedge design was planned with a total of 39 clinical units to be engaged (Table 1) [36,37]. SOS was to roll out in three Waves: Wave 1 (months 1–12), the six EDs located across our four hospitals; Wave 2 (months 13–30), the 25 inpatient medical and psychiatric units; and, Wave 3 (months 31–48), eight large primary care and obstetrics/gynecology (OB/GYN) clinics.

Within each Wave, the individual units were to be randomly assigned to a cohort, with each cohort randomly assigned to start dates three months apart. Each cohort was to progress through three phases: Preparation (3 months), Implementation (12 months), and Sustainment (12 months). During Preparation, the System Hub was to work with Local Hubs and Spokes to prepare clinical protocols tailored to each clinical unit. These protocols would be implemented and refined during Implementation and further improved and maintained during Sustainment.

Study population. As depicted in Table 1, all patients 12 years and older seen in any of the 39 units during the monitoring period before the

unit’s Implementation phase were considered Control patients. All such patients presenting during or after the unit began Implementation were Intervention patients. As new units entered the study, any patient previously classified as a Control or Intervention maintained that initial designation and was not reclassified. For example, when we estimated that 60,000 patients would newly join the study in Wave 2 (inpatient), this was an estimate of the number admitted to a study inpatient unit who had not been previously considered study patients (Control or Intervention) due to an ED visit.

Based on historical data, we estimated that the total number of study patients accrued over four years would be ~310,000 (Table 1), about 50% female, 14% Latinx, and 29% from under-represented minorities, including Latinx of any race, and 15% children 12–18 years old.

Data collection and measurement. We intended to use EHR data, including clinical and claims data, for both CQI and research. While data directly downloaded from structured EHR fields (i.e., program-derived data) would be our main source, we planned to validate the EHR data by (1) structured manual reviews by trained research staff, and (2) direct interviews with patients by research staff, referred to as fidelity checks. EHR data was supplemented by Massachusetts (MA) vital statistics for tracking deaths and suicides. Patient outcomes, intervention targets, potential mechanisms and moderators, fidelity, healthcare services utilization, and costs were to be measured as described below.

Patient outcomes. We envisioned two primary patient outcomes, suicide risk identification and the suicide-related composite, identified from EHR program-derived data and MA vital statistics. *Suicide risk identification* was defined as patient endorsement of active ideation in the past two weeks OR report of a lifetime suicide attempt on the risk screening tool. *Suicide composite outcome* was defined as probable death by suicide OR a suicide-related acute care episode. Suicide-related acute care episodes are ED visits or hospitalizations for a suicide attempt or suicidal ideation. There were three potential sources for determining suicide risk identification and suicide composite outcomes: (1) ICD codes associated with suicide ideation, attempts, or acute care episodes resulting in death (detailed in appendix 1); (2) program-derived data obtained from the structured suicide screening instrument included in the EHR as part of SOS; and (3) MA vital statistics (particularly for data on probable death by suicide).

Intervention targets (clinician behaviors). We mapped each of the targeted clinical performance elements (see above) to the relevant EHR fields to guide EHR data download and standardized manual chart abstractions. We tracked standardized suicide risk screening (completion of suicide screen), safety planning (collaborative written document developed with the patient), means restriction counseling, and care-transition facilitation (including post-acute telephone follow-ups within 24-h of discharge).

Moderators, mediators, mechanisms of action would be assessed from

A3 problem-solving	A pragmatic problem-solving approach and communication plan for creating sustainable change. The iterative steps of the A3 process are: 1. Define the problem statement or gap in performance (i.e., not screening, not providing safety planning) 2. Understand and map the current process 3. Determine the root causes of the problem 4. Develop the actions to address root causes and plan the interventions 5. Implement the plan 6. Collect follow-up data on the outcome of the new process and compare it against pre-specified targets and if needed, act with newly designed actions Steps 2-6 are usually carried out as a series of cycles until the desired target is met.
Kaizen events	A focused, short-term event that brings together stakeholders to support the execution of A3s.
Results board	Visual display of goal progress for ongoing departmental and personnel performance improvement. The A3 is posted next to the results board as a communication tool and as a way to demystify the plan to the organization.

Fig. 3. Definitions of Core Lean Tools to be Implemented During SOS.

Table 1
Stepped wedge design of clinical units entering Control (C) vs. Intervention (I).

Time (mos)	Wave 1 (12 months)				Wave 2 (18 Months)						Wave 3 (18 Months)				F/U (6 Months)		
	1	4	7	10	13-	16-	19-	22-	25-	28	31-	34-	37-	40-		43	49
	-	-	-	-	13-	16-	19-	22-	25-	-	31-	34-	37-	40-	-	-51	
	3	6	9	12	15	18	21	24	27	30	33	36	39	42	46		
Unique new patients : n*	100,000				60,000						150,000						
EDs																	
2	C	I	I	I	I	I	I	I	I	I	I	I	I	I	I	I	
2 Units	C	C	I	I	I	I	I	I	I	I	I	I	I	I	I	I	
2 Units	C	C	C	I	I	I	I	I	I	I	I	I	I	I	I	I	
INPATIENT UNITS (n= 25)																	
5 Units					C	I	I	I	I	I	I	I	I	I	I	I	
5 Units					C	C	I	I	I	I	I	I	I	I	I	I	
5 Units					C	C	C	I	I	I	I	I	I	I	I	I	
5 Units					C	C	C	C	I	I	I	I	I	I	I	I	
5 Units					C	C	C	C	C	I	I	I	I	I	I	I	
OUTPATIENT CLINICS (n=8)																	
2											C	I	I	I	I	I	
2											C	C	I	I	I	I	
2											C	C	C	I	I	I	
2											C	C	C	C	I	I	
Major disruptions to planned study design***																	
EHR Vendor Transition																	
New Joint Commission																	
Leadership Changes																	
Macro System Changes																	
COVID-19																	

four sources: (1) *Clinician surveys* targeting knowledge, attitudes, efficacy, and practice related to suicide and perceived health system and departmental leadership support for SOS were measured at Months 1, 30, 40. (2) *Lean team observations* measured enabling attributes such as embedded mental health clinicians and clinician-level performance feedback, using structured Lean evaluations. (3) *EHR data* on outpatient behavioral treatment (e.g., time of engagement), clinically relevant covariates such as comorbidity and health insurance status and type. (4) *Geographic Information Systems (GIS)* to provide data on markers of neighborhood deprivation.

Fidelity to the Zero Suicide Model was tracked throughout the study. *Screening fidelity* was measured by a research assistant completing a brief interview with a random sample of 20 patients per unit whose records indicated absence of suicidal ideation/behavior during the first and last months of implementation in that unit. They identified whether the screening was performed (vs simply documented). The Visual CARE 5Q was administered to measure patients' satisfaction with care [38]. A sample of safety plans were to be rated using a structured quality rating tool [39]. to assess *fidelity to the SPI* and lethal means safety counseling intervention (Step 6 of the Safety Plan). *Lean fidelity* was measured by documenting whether the Spokes established a local team and defined the problem, gathered background data and process observations, used at least three Lean techniques, and applied PDSA cycles adapted as needed.

Healthcare costs, specifically suicide-related direct healthcare services utilization and costs associated with SOS implementation were to be measured. Costs are estimated from health system reimbursement rates (average DRG rates) using cost-to-charge ratios.

2.1. Data analytic plan

The choice of a stepped wedge design rather than a simple cluster-randomized trial was based on the need to rapidly implement SOS across the entire healthcare system while still maintaining randomization for causal inference. With a stepped wedge design, each study unit would receive the intervention within the study Wave to which it was assigned. Unfortunately, the disruptions described in the next section interrupted this approach. We are now analyzing our data using an interrupted time series approach, still appropriate for testing some of SOS's main hypotheses.

3. Study disruptions and adaptations

Above, we described our planned intervention and research methods. Now, we describe the obstacles faced, and the necessary adaptations applied, throughout the study period. Five major disruptions impacted implementation and research methods: (1) the healthcare system switched EHR vendors midway through Wave 1; (2) The Joint Commission, which sets safety standards for hospitals, changed its suicide risk mitigation standard; (3) health system leadership changed; (4) macro health system changes resulted in acquisitions, unit closures, and new units; and (5) COVID-19 emerged. The lower section of Table 1 places the duration of these disruptions in the timeline of the planned design. Table 2 summarizes the disruptions' consequences, including major lessons learned.

EHR vendor switch. In October 2017, 19 months after the start of SOS, the health system moved from dozens of setting-specific EHRs to an

Table 2
Major disruptions of SOS implementation plan, harms and benefits, and lessons learned.

Disruption	Harms	Benefits & Adaptations	Lessons learned
EHR Switch	System resources stretched with less leader and staff engagement	New tools allowed for better cross-unit communication and automation of discharge orders; Study team worked directly with EHR programmers	Embrace opportunity to recast care using new tools, but beware of concomitant disruption of clinical activities
	Difficulty using new tools	Better reporting for Quality Improvement	Care can be enhanced across settings and efficiency improved with enterprise (cross setting) EHR
	Data format and structure was different across EHRs	Harmonization of data across EHRs was necessary	Data harmonization across EHRs is both crucial and time-consuming
Standards change	Made experimental design unfeasible. Moved from stepped-wedge to interrupted time series	Goals of Zero Suicide were embraced much more fervently	Sentinel adverse events can be major engines of beneficial change but may require changes in research methods, especially if the changes impact all units or entities
Leadership changes	Withdrawal of resources	Potential new champions	Negotiate study support and resources that is not dependent on specific leaders early on
	Delays in implementation	Opportunity to review and modify work-flow with leadership and staff	Be patient and recognize efforts may lag while leadership is in flux but be persistent
	May act as a secular event or interruption that impacts all units		May require statistical modeling as a covariate or interruption
Macro health system changes	Additional burden of new units	Addition of new implementation sites and related data collection	Health care is dynamic and volatile; when possible, build in flexibility in implementation plans and study design
	Removal of original units		Adjust data acquisition, validation according to new demands
	Changes in data sources, quality Changes in sample size		
COVID-19	Delays in implementation as priorities shift	None	Be prepared to model major external disruptions in the analyses of

Table 2 (continued)

Disruption	Harms	Benefits & Adaptations	Lessons learned
			any implementation study
	May act as a secular event or interruption that impacts all units		
	Leads to staff burnout, which leads to staff shortages and resistance to new efforts		

enterprise solution that cut across the entire organization. This change both positively and negatively affected our implementation efforts. Positively, the new EHR was enhanced by the SOS team with suicide-related tools before going live. This improved the health system’s ability to standardize instruments, prompt best-practices through clinical decision support tools, and facilitate cross-setting communication of risk. For example, we built a personalized safety planning template that could “follow the patient” across settings: a plan created in the ED could be accessed, used, and edited by inpatient providers. This ability to share tools and alerts across settings helped solve some trenchant communication gaps, such as a patient’s suicide risk identification in one setting not being communicated to another setting. Finally, the new EHR provided data reporting absent from our previous EHR, with improved CQI support by quantifying metrics across the system. The new EHR led to a unique state of flux and customization, allowing us to propose novel tools that might have been difficult to implement under more routine circumstances.

The negative effects centered around the new EHR implementation overwhelming all other competing health system priorities, with decreased engagement of leaders and frontline staff in SOS preparation and implementation. Further, while the new EHR had novel reporting features, they were difficult to use, and did not easily allow clinician-level performance reporting. This resulted in a net improvement of macro level metric monitoring but reduced micro level auditing.

Important generalizable lessons were learned for implementation efforts. EHRs, now ubiquitous, can help standardize measurement, support improved care protocols, and improve communication across settings, especially when deployed enterprise-wide. A change in EHR provides tremendous opportunities to recast care, but may also disrupt improvement efforts, which needs to be incorporated into projected timelines. EHR-related tools work better if they are used across settings and programmed to allow “carry forward” information, thus improving cross-setting communication and efficiency.

Regarding research methods, the EHR change introduced a secular event, or an interruption in the terminology of interrupted time series designs, that impacted all units across the system, which disrupts the carefully laid plan of exposure in timed Waves or steps. It also substantially impaired our ability to interpret our EHR data. Comparisons that would appear to be straightforward, such as validating our primary outcome definitions, have been complicated by variables in the new EHR not meaning exactly the same as their homonyms in the previous software. When designing a longitudinal study over a period bridging different EHRs, the challenges of harmonizing such data should not be underestimated.

Changes in the Joint Commission’s standards. For the first 1.5 years, our implementation efforts rolled out as originally proposed. Wave 1 (six EDs) and Wave 2 (five inpatient units) proceeded based on our cohort steps depicted in [Table 1](#) and each cohort progressed through Preparation, Implementation, and Sustainment phases. However, as we were working with inpatient Cohort 2 within Wave 2, one of our system hospitals was cited as non-compliant with the recently updated suicide

prevention standard published by The Joint Commission. This was a major disrupter, as the organization's leadership immediately enforced several changes across the system of hospitals in-order to maintain compliance with the new standards.

This change had a profound impact on SOS. Rather than working with individual units and fashioning tailored approaches through Spokes, a large, system-wide team established by leadership and built policies and protocols to be enforced across the entire healthcare system and for all acute care units. Further, leadership buy-in and tangible support of suicide prevention expanded dramatically, with substantial resources devoted towards the effort. Edits and optimization of the EHR tools the study team had worked to build proceeded slowly prior to the Joint Commission visit; after, however, they were escalated to the top of the worklist. Before the visit, training clinicians on the new protocols was weak and inconsistent; after the visit, intensive training of clinicians across the entire health system became a priority.

Thus, major changes in standards or regulation can have unanticipated positive and negative consequences. This change transformed our efforts in a positive way by providing a "burning platform" to mobilize leadership and resources. This is common in healthcare; change happens at a glacial pace, unless there is a major event inspired by regulatory, accreditation, or financial incentives, which catalyzes rapid change. It is impossible, however, to manufacture such a motivator like a Joint Commission standard change.

The system-wide changes induced by the Joint Commission also affected our research design and statistical plan. Rather than rolling out efforts using the stepped wedge design, our program was implemented across all remaining Control and Intervention acute care units at once. This led to us changing the design to an interrupted time series study. Time series designs, while commonly used in implementation studies, are somewhat weaker in terms of establishing causality. Fortunately, we were able to continue to collect data that would be amenable to testing our hypotheses, albeit with a different analytic approach. There is also the added opportunity to study the effect of such major disruptions via a "natural experiment."

Leadership changes. Throughout our implementation, changes in leadership at many levels influenced our capabilities to achieve system change. The changes weakened the resources to implement our efforts and created a power vacuum that, when combined with COVID-19 (below), completely halted progress with rolling out Wave 3 (ambulatory care). New leaders eventually were hired, but it was over a year after the originally planned implementation; COVID-19 had completely rearranged the health system's priorities, impeding implementation efforts.

The lesson learned is clear. High level executive sponsorship that results in tangible support, such as CQI resources and protected clinician time for training, is required for effecting and sustaining system-based change; without this, efforts get under-resourced, deprioritized, and stagnate or fail. Careful attention to laying the groundwork and describing the needed tangible supports, including, if needed, creating budgets and an estimate of the return on investment, are advisable. Inevitably, changes in leadership will occur; anticipating these changes is difficult but putting into place a team that can continue across leadership changes, formalizing the efforts into written policies and procedures, and upskilling team members, can offer some protection.

The leadership changes convey important implications for research methods. First, as with other changes above, they introduce a secular event or possible "interruption" that can lead to their own influence on the implementation efforts. They may need to be modeled in any resulting analyses or tested for their confounding effect. Leaders' mandates can impact the timeline of the study, pushing all efforts, deprioritizing data collection and analyses, and challenging achievement of milestones and completion of the study during the allotted time.

Macro health system changes. As a result of an acquisition, we had to add an entire hospital to our implementation and data collection. The new hospital has a large ED and several medical units and is a

community-based hospital with no academic mission. These factors combined to create unanticipated burden impeding implementation. It not only expanded the scope of the project by 20%, but it also slowed everything down by adding a new set of leaders, administration, existing policies, and culture with which to contend.

The other big change occurred in the outpatient setting (Wave 3). One of our original eight primary care clinics closed, but we added one large primary care clinic and a large outpatient mental health clinic. The net effect of these changes was to further increase the scope of the outpatient effort by about 25% and added complexity in the team representation and administrative bureaucracy.

The primary implementation lesson from this is that healthcare is dynamic and volatile. Teams should prepare for changes, such as dropping and adding new units, and adapt to these changes in as efficient a manner as possible. If units or clinics are added or dropped, the overall patient and provider populations being examined will likely change. The primary research method lesson is that the source data may change if clinics are added or dropped. It can also fundamentally affect the sample size, both at the unit and the patient level, which can impact the stability of findings, positively or negatively.

COVID-19. During our preparation for Wave 3, COVID-19 became prevalent in our region, profoundly altering the provision of healthcare at all levels and creating an over-riding priority focus. In the short term, all attention diverted to COVID-19, de-emphasizing suicide prevention and most other performance improvement. Much outpatient behavioral healthcare adopted telehealth, disrupting usual screening processes, while ED visits fell as patients avoided seeking acute care. In the longer term, COVID-19 has led to burned out clinicians and massive staff turnover, with resistance to new activities. COVID-19, combined with the power vacuum described above, halted our ambulatory roll out efforts for more than a year, and the renewed effort was very slow while the health system recuperated. COVID-19 will likely continue to exert significant pressure on health systems for some time, further subordinating our efforts to improve suicide-related care.

External events like pandemics cannot be predicted. Implementation efforts may be delayed or cancelled; implementation teams and health systems will have to evaluate and prioritize these efforts, perhaps resulting in less intensive and or incomplete implementation. For example, while one may want to provide intensive skills-based training for screening and safety planning, one may have to rely more heavily on easier to access and less time-consuming virtual trainings.

Research-wise, pandemics and other events that impact entire societies may need to be modeled. For example, COVID has significantly impacted suicide-related presentations to EDs. It appears that there was a brief decrease in ED visits in March 2019, followed by an increase ED visits in the latter part of 2020 compared to the same time period in 2019 [4]. Because part of our primary outcome is derived from suicide-related acute care, this can have a major confounding influence on our findings.

In summary, as is not unusual in implementation research, several large-scale and unexpected events affected our detailed plans for both the clinical implementation of Zero Suicide and our study of this implementation. Some of the encountered obstacles were partially predictable, such as leadership changes and the difficulty in enhancing care transitions, while others were quite unpredictable, such as a major pandemic. We have learned multiple lessons from SOS, which we hope will contribute to the collective experience and learning processes of the implementation research community. Specific analyses of our very rich data will contribute generalizable knowledge despite the numerous barriers encountered.

Funding

The work was supported by the National Institute of Mental Health [Award Number R01MH112138]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National

Institutes of Health.

Collaborators

This paper was written on behalf of the SOS Investigator Team.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix 1. ICD codes used to identify a suicide-related acute care episode, suicidal ideation, or suicide attempt

ICD-10-CM codes for suicide attempt and intentional self-harm.

- Intentional self-injury by various means: X71.2xx through X83.8xx
- Poisoning by various means: T36.1x2 through T50.Z92
- Toxic effects by various means: T51.0x2 through T65.92x
- Asphyxiation by various means, intentional self-harm: T71.112 through T71.232
- Suicide attempt T14.91
- Suicidal ideation R45.851

References

- [1] Centers for Disease Control and Prevention CDC, 10 leading causes of death by age group, *United States - 2014*. Retrieved from: https://www.cdc.gov/injury/wisqars/pdf/leading_causes_of_death_by_age_group_2014-a.pdf, 2014.
- [2] F.B. Ahmad, R.N. Anderson, The leading causes of death in the US for 2020, *JAMA* 325 (18) (2021) 1829–1830.
- [3] S.A. Ting, et al., Trends in US emergency department visits for attempted suicide and self-inflicted injury, 1993–2008, *Gen. Hosp. Psychiatr.* 34 (5) (2012) 557–565.
- [4] K.M. Holland, et al., Trends in US emergency department visits for mental health, overdose, and violence outcomes before and during the COVID-19 pandemic, *JAMA Psychiatr.* 78 (4) (2021) 372–379.
- [5] M.H. Allen, et al., Screening for suicidal ideation and attempts among emergency department medical patients: instrument and results from the Psychiatric Emergency Research Collaboration, *Suicide Life-Threatening Behav.* 43 (3) (2013) 313–323.
- [6] M.A. Ilgen, et al., Recent suicidal ideation among patients in an inner city emergency department, *Suicide Life-Threatening Behav.* 39 (5) (2009) 508–517.
- [7] E.D. Boudreaux, et al., Improving suicide risk screening and detection in the emergency department, *Am. J. Prev. Med.* 50 (4) (2016) 445–453.
- [8] C.A. Claassen, G.L. Larkin, Occult suicidality in an emergency department population, *Br. J. Psychiatry* 186 (2005) 352–353.
- [9] E.D. Boudreaux, et al., The patient safety screener: validation of a brief suicide risk screener for emergency department settings, *Arch. Suicide Res.* 19 (2) (2015) 151–160.
- [10] M. Olsson, S.C. Marcus, J.A. Bridge, Emergency department recognition of mental disorders and short-term outcome of deliberate self-harm, *Am. J. Psychiatr.* 170 (12) (2013) 1442–1450.
- [11] M. Olsson, S.C. Marcus, J.A. Bridge, Emergency treatment of deliberate self-harm, *Arch. Gen. Psychiatr.* 69 (1) (2012) 80–88.
- [12] National Action Alliance for Suicide Prevention, Clinical care & intervention task force. Suicide care in systems framework, Retrieved from, <http://actionallianceforsuicideprevention.org/sites/actionallianceforsuicideprevention.org/files/taskforces/ClinicalCareInterventionReport.pdf>, 2011.
- [13] J. Pearson, N.I.o.M.H. (NIMH), in: *Applied Research towards Zero Suicide Healthcare Systems*, 2015. Washington, DC.
- [14] Suicide Prevention Resource Center, ZEROsuicide in health and behavioral health care, Available from: <http://zerosuicide.edc.org/>, 2020.
- [15] C.E. Coffey, Building a system of perfect depression care in behavioral health, *Joint Comm. J. Qual. Patient Saf.* 33 (4) (2007) 193–199.
- [16] M.J. Coffey, C.E. Coffey, B.K. Ahmedani, Suicide in a health maintenance organization population, *JAMA Psychiatr.* 72 (3) (2015) 294–296.
- [17] C.A. Hoffmire, J.E. Kemp, R.M. Bossarte, Changes in suicide mortality for veterans and nonveterans by gender and history of VHA service use, 2000–2010, *Psychiatr. Serv.* 66 (9) (2015) 959–965.
- [18] K. Kryszinska, et al., Best strategies for reducing the suicide rate in Australia, *Aust. N. Z. J. Psychiatr.* 50 (2) (2016) 115–118.
- [19] C.D. Labouliere, et al., Zero Suicide® - a model for reducing suicide in United States behavioral healthcare, *Suicidologi* 23 (1) (2018) 22–30.
- [20] K. Turner, et al., Implementing a systems approach to suicide prevention in a mental health service using the Zero Suicide Framework, *Aust. N. Z. J. Psychiatr.* 55 (3) (2021) 241–253.
- [21] I. Education Development Center, ZEROsuicide data elements worksheet. http://ze.rosuicide.edc.org/sites/default/files/ZS%20Data%20Elements%20Worksheet.TS_.pdf, 2015.
- [22] C.M. van der Feltz-Cornelis, et al., Best practice elements of multilevel suicide prevention strategies: a review of systematic reviews, *Crisis* 32 (6) (2011) 319–333.
- [23] The Joint Commission, The suicide prevention portal. <https://www.jointcommission.org/resources/patient-safety-topics/suicide-prevention/>, 2020.
- [24] E. Boudreaux, I. Miller, C. Camargo, Emergency department safety assessment and follow-up evaluation (ED-SAFE) trial, National Institute of Mental Health. U01MH088278, <https://reporter.nih.gov/project-details/8484443>, 2009.
- [25] E.D. Boudreaux, et al., Emergency department safety assessment and follow-up evaluation 2: an implementation trial to improve suicide prevention, *Contemp. Clin. Trials* 95 (2020), 106075.
- [26] B.J. Powell, R.S. Beidas, Advancing implementation research and practice in behavioral health systems, *Adm Policy Ment Health* 43 (6) (2016) 825–833.
- [27] A.A. Fagan, et al., Scaling up evidence-based interventions in US public systems to prevent behavioral health problems: challenges and opportunities, *Prev. Sci.* 20 (8) (2019) 1147–1168.
- [28] K.L. Carmen, K. Paez, J. Stephens, Improving Care Delivery through Lean: Implementation Case Studies, American Institutes for Research, Urban Institute, Mayo Clinic, Rockville, MD, 2014.
- [29] H. Andersen, K.A. Rovik, T. Ingebrigtsen, Lean thinking in hospitals: is there a cure for the absence of evidence? A systematic review of reviews, *BMJ Open* 4 (1) (2014), e003873.
- [30] J.S. Toussaint, L.L. Berry, The promise of Lean in health care, *Mayo Clin. Proc.* 88 (1) (2013) 74–82.
- [31] H.C. Kaplan, et al., An exploratory analysis of the model for understanding success in quality, *Health Care Manag. Rev.* 38 (4) (2013) 325–338.
- [32] H.C. Kaplan, et al., The Model for Understanding Success in Quality (MUSIQ): building a theory of context in healthcare quality improvement, *BMJ Qual. Saf.* 21 (1) (2012) 13–20.
- [33] H.C. Kaplan, et al., The influence of context on quality improvement success in health care: a systematic review of the literature, *Milbank Q.* 88 (4) (2010) 500–559.
- [34] A.C. Feldstein, R.E. Glasgow, A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice, *Joint Comm. J. Qual. Patient Saf.* 34 (4) (2008) 228–243.
- [35] L.J. Damschroder, et al., Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science, *Implement. Sci.* 4 (2009) 50.
- [36] C.A. Brown, R.J. Lilford, The stepped wedge trial design: a systematic review, *BMC Med. Res. Methodol.* 6 (2006) 54.
- [37] M.A. Hussey, J.P. Hughes, Design and analysis of stepped wedge cluster randomized trials, *Contemp. Clin. Trials* 28 (2007) 182–191.
- [38] M.A. Place, et al., A preliminary evaluation of the Visual CARE Measure for use by Allied Health Professionals with children and their parents, *J. Child Health Care* 20 (1) (2016) 55–67.
- [39] G. Brown, B. Stanley, Coding Manual for Rating of Safety Plans in Medical Records, University of Pennsylvania and Columbia University, 2013.