



ELSEVIER

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial

Emergency department safety assessment and follow-up evaluation 2: An implementation trial to improve suicide prevention

Edwin D. Boudreaux^{a,*}, Brianna L. Haskins^b, Celine Larkin^b, Lori Pelletier^c, Sharon A. Johnson^d, Barbara Stanley^e, Gregory Brown^f, Kristin Mattocks^g, Yunsheng Ma^g^a Departments of Emergency Medicine, Psychiatry, and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, USA^b Department of Emergency Medicine, University of Massachusetts Medical School, Worcester, MA, USA^c Department of Pediatrics, University of Connecticut School of Medicine, Hartford, CT, USA^d Foiesie Business School, Worcester Polytechnic Institute, Worcester, MA, USA^e Department of Psychiatry, Columbia University, New York, NY, USA^f Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA^g Department of Population and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, USA

ARTICLE INFO

Keywords:

Suicide
Suicide prevention
Mental health
Quality improvement
Implementation science

ABSTRACT

Background: Emergency departments (EDs) are important for preventing suicide. Historically, many patients with suicide risk are not detected during routine clinical care, and those who are often do not receive suicide-specific intervention. The original Emergency Department Safety Assessment and Follow-up Evaluation (ED-SAFE 1) study examined the implementation of universal suicide risk screening and a multi-component ED-initiated suicide prevention intervention.

Purpose: The ED-SAFE 2 aims to study the impact of using a continuous quality improvement approach (CQI) to improve suicide related care, with a focus on improving universal suicide risk screening in adult ED patients and evaluating implementation of a new brief intervention called the Safety Planning Intervention (SPI) into routine clinical practice. CQI is a quality management process that uses data and collaboration to drive incremental, iterative improvements. The SPI is a personalized approach that focuses on early identification of warning signs and execution of systematic steps to manage suicidal thoughts. ED-SAFE 2 will provide data on the effectiveness of CQI procedures in improving suicide-related care processes, as well as the impact of these improvements on reducing suicide-related outcomes.

Methods: Using a stepped wedge design, eight EDs collected data cross three study phases: Baseline (retrospective), Implementation (12 months), and Maintenance (12 months). Lean methods, a specific approach to pursuing CQI which focuses on increasing value and eliminating waste, were used to evaluate and improve suicide-related care.

Conclusions: The results will build upon the success of the ED-SAFE 1 and will have a broad public health impact through promoting better suicide-related care processes and improved suicide prevention.

1. Background

Suicide is the 10th leading cause of death in the US [1]. Modern prevention efforts have hailed emergency departments (EDs) as important settings for preventing suicide, where both manifest risk,

(individuals presenting with suicidal behavior), and latent risk, (individuals presenting with medical complaints but who have recently thought about suicide), are prevalent [2–6]. Historically, many at-risk patients are not detected during routine care, and, those who are, often do not receive suicide-specific intervention [7–8]. To help address this,

Abbreviations: emergency department, (ED); Emergency Department Safety Assessment and Follow-Up Evaluation, (ED-SAFE); continuous quality improvement, (CQI); safety planning intervention, (SPI); Veterans Health Administration, (VHA); Practical, Robust Implementation and Sustainability Model, (PRISM); Consolidated Framework for Implementation Research, (CFIR); Plan-Do-Study-Act, (PDSA); Patient Safety Screener, (PSS-3); Safety Plan Intervention Rating Scale, (SPIRS); electronic health record, (EHR); National Death Index, (NDI); generalized estimating equation, (GEE); research assistant, (RA)

* Corresponding author at: 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), Brianna.haskins@umassmed.edu (B.L. Haskins), Celine.larkin@umassmed.edu (C. Larkin), lpelletier01@connecticutchildrens.org (L. Pelletier), sharon@WPI.EDU (S.A. Johnson), bhs2@cumc.columbia.edu (B. Stanley), gregbrown@penmedicine.upenn.edu (G. Brown), Kristin.mattocks@umassmed.edu (K. Mattocks), yunsheng.ma@umassmed.edu (Y. Ma).

<https://doi.org/10.1016/j.cct.2020.106075>

Received 19 February 2020; Received in revised form 12 June 2020; Accepted 15 June 2020

Available online 19 June 2020

1551-7144/ © 2020 Elsevier Inc. All rights reserved.

the original Emergency Department Safety Assessment and Follow-up Evaluation (ED-SAFE 1) study examined the implementation of universal suicide risk screening and a multi-component ED-initiated suicide prevention intervention [9]. The ED-SAFE 1 concluded that: [1] using common continuous quality improvement (CQI) strategies, it was feasible to increase suicide screening to reach, on average, 85% of the adult ED population across eight diverse EDs; [2] such an increase in screening nearly doubled risk detection, rising from 2.9% to 5.7% of adults [7]; and, [3] implementing a multi-component intervention, including post-visit telephone calls, led to a 30% relative reduction in suicidal behavior compared to treatment as usual or screening alone [10].

This paper describes the methodology used in the ED-SAFE 2 study, which builds upon the ED-SAFE 1 to further evaluate, using a CQI strategy called Lean, improvement of suicide-related care processes. The current manuscript provides the necessary detail for judging the quality and rigor of the ED-SAFE 2 methods, as a foundation for forthcoming results papers as well as a guide for future studies. While all suicide-related care was targeted, particular focus was on the following aims: [1] continued improvement and long-term sustainability of universal suicide risk screening in adult ED patients, and [2] to evaluate implementation of a new brief intervention called the Safety Planning Intervention (SPI) into routine clinical practice. The SPI was chosen because of its empirical support and potential for translation into clinical care [2]; SPI is a best practice recommended by the Joint Commission [11] and the Suicide Prevention Resource Center.

A recent study evaluated the efficacy of SPI at nine Veterans Health Administration (VHA) EDs [11]. The SPI was administered to Veterans with a suicide-related concern and for whom inpatient hospitalization was not clinically indicated, and included telephone contact following discharge. In this study [$n = 1650$], the SPI plus follow-up was associated with about 50% fewer suicidal behaviors over a 6-month follow-up and more than doubled the odds of engaging in outpatient behavioral health care, reinforcing the utility of the SPI as an effective prevention strategy in EDs.

ED-SAFE 2 will provide essential data on the effectiveness of CQI in improving suicide-related care, including screening and the SPI, and in understanding how these improvements impact suicide-related outcomes. It will extend the findings of Stanley et al., 2018 [12] with focus on implementation in civilian EDs without the post-visit call. In contrast to the VHA study, which employed study staff to provide the intervention, ED-SAFE 2 will use clinical staff. Healthcare services researchers have advanced understanding of how to implement evidence-based practices into routine clinical practice; however, as Chambers and colleagues [13] and others [14–16] have stated, less is known about the long-term maintenance of these improvements. The ED-SAFE-2 will leverage study data to evaluate both implementation and sustainability of suicide-related process improvements, including examining organizational factors that impact success.

2. Study design and procedures

The study used a stepped wedge design [17], which is well adapted to effectiveness trials and implementation studies where randomization by patient is impractical. Eight sites were split into two groups based on a median split of high vs. low screening performance in ED-SAFE 1. Each high performing site was randomly paired with a low performing site to create a two-site cohort that prevented confounding of starting order and performance which might occur by random assignment without this balancing (i.e., all the high performing sites could have been randomly assigned to start earlier or later by chance). The cohorts were then randomly assigned to a start date (wedge) separated by two months. Standardized data collection procedures were used across all three phases (see Section 2.4 for details).

Each of the eight EDs collected data across three study phases: Baseline, Implementation, and Maintenance. Baseline established the

current practices and patient outcomes and acted as the control condition. During the Implementation and Maintenance phase, a CQI approach called Lean was used to foster implementation, improvement, and sustainability of screening, the SPI, and other suicide-related process improvements identified by each site through their CQI process. The intervention targets were at the clinician/behavioral level and comprised clinician suicide risk screening and, for those with suicide risk but discharged from the ED, personalized safety planning, i.e., the SPI. The intervention consisted of refresher training on universal screening, new training on the SPI, and using CQI teams (called “Lean Teams”) to identify and remediate care gaps, and to monitor, improve, and sustain these efforts. Sites’ Lean Teams were established by each site PI. Each team received in-person Lean training early in Implementation, delivered by one of two co-investigators who are experts in Lean and CQI. The Lean Teams were self-led, multidisciplinary, and met locally on a monthly or quarterly basis to build and update CQI tools (including A3s and run-charts). Site-level CQI activities were supported by, and reported activity to, the coordinating center on a monthly basis throughout Implementation and Maintenance. The hypothesized mechanisms of action of our intervention included improving ED organizational characteristics, such as leadership support and staff attitudes, and building enabling infrastructure, such as integration of tools into the electronic health record (EHR), that support suicide-related care processes such as screening and safety planning. These mechanisms were chosen because they were noted as important during qualitative interviews in the original ED-SAFE 1 study; they have broad support in the CQI literature [18–19]; they are intrinsic to Lean principles; and they play a prominent role in many implementation science models. While they are most clearly conceptualized in the Practical, Robust Implementation and Sustainability Model (PRISM) [20], they also appear in other models like the Consolidated Framework for Implementation Research (CFIR) [21]. Initial competence and longitudinal fidelity in screening, the SPI, and Lean were determined. The patient outcomes were suicide risk detection and a suicide composite outcome comprised of suicide or a suicide-related acute health-care visit within the 6 months post-ED visit.

2.1. Study phases

Each site progressed through the following phases:

2.1.1. Phase 1: baseline

The 12 months prior to the preparation period comprised Baseline. This represented a period after enrollment into the original ED-SAFE 1 study and before the onset of the new preparation activities for ED-SAFE 2. Charts during this period were retrospectively reviewed to [1] examine screening rates post-ED-SAFE 1 under naturalistic conditions and [2] to establish a baseline comparator for screening, safety planning, and other care processes against which to compare performance during ED-SAFE 2 Implementation.

2.1.2. Phase 2: implementation

During the 12 months of Implementation, each site used a team that had been trained in Lean principles and tools by a Master Lean Trainer to identify and implement process improvements related to suicide care, including but not limited to universal screening and the SPI as required components. Each local Lean Team oversaw and monitored implementation of all targeted suicide related care processes into routine clinical care. During the three months prior to a site’s “go live” date, site’s clinicians were trained to deliver the SPI by a Master SPI Trainer. These clinicians were most often mental health counselors and social workers; sites chose who should conduct safety plans based on the site’s determination of best-fit with the SPI intervention and workflow.

2.1.3. Phase 3: maintenance

The Implementation phase was followed by a 12-month Maintenance phase during which the site worked to maintain or further improve upon the gains made during Implementation.

2.2. Intervention components

2.2.1. Lean

Lean CQI techniques were used to evaluate and improve performance of suicide-related care. The CQI team (Lean Team) at each site had representation from ED physicians, nurses, and mental health clinicians, and was coached by one of the study co-investigators with Lean expertise. While the site teams were encouraged to improve any suicide-related care processes, they were each tasked with [1] maximizing universal suicide screening and [2] implementing safety planning for those with suicide risk who were being discharged.

Lean includes a diverse range of tools to foster implementation. Because the focus in ED-SAFE 2 was to implement a departmental change in care processes and clinician behavior, three tools that align well with this purpose were used to define the Lean intervention. Teams developed an overall plan for solving an identified problem called an A3 plan, which creates a structure for understanding current conditions, identifying solutions and executing them through continuous improvement cycles. The A3 plan included one or more Plan-Do-Study-Act (PDSA) cycles, where generated solutions were tested and results were measured and discussed to inform additional actions for improvement. Result boards were used to inform, educate and celebrate the improvement activities and progress with all clinicians. Sites were encouraged to employ other Lean tools in addition to these, describing their use and providing supporting documents.

The site Lean Team members were trained by study co-investigators, who were Lean Master Trainers, to ensure a common level of knowledge about Lean philosophy and techniques. The training included didactics and case studies and was heavily focused on employing Lean principles and tools with the site's team in real time. In particular, the team initiated a broad A3 plan based on the goal of using screening and safety planning to meet target thresholds, which included creating a process map to describe how the screening and safety planning would be integrated at the site. After the training, sites used PDSA cycles to test the process, collected data relevant to their targeted process improvement, and reviewed results to generate follow-up improvement cycles, which were documented in their A3 plans. During Implementation, study co-investigators held monthly coaching calls with the Lean Team at each site to root cause issues and support improvement cycles.

Competence in Lean was established by requiring that Lean Team members score above an 80% threshold level on a competency survey assessing knowledge of Lean principles and practices. Team competence in Lean was further assured by comparing the output of the team's first improvement cycle against Lean standards, including: [1] an updated A3 documenting the overall problem and goal, [2] results depicting initial implementation of a change cycle, and [3] identification of at least one issue the team would address in a future improvement cycle.

2.2.2. Universal screening

"Universal screening" refers to systematic screening of all patients for a health outcome or risk factor. Universal screening for suicide risk in the emergency department entails screening every patient regardless of presenting complaint and is a practice recommended by the Joint Commission [11] due to an overrepresentation of suicide risk in emergency department patients [22–24]. During ED-SAFE 1, all eight sites had adopted clinical protocols for universal suicide risk screening using the Patient Safety Screener [7]. The Patient Safety Screener (PSS-3) consists of three items, assessing depressed mood and suicidal ideation in the past two weeks and lifetime suicide attempt [24]. A positive screen was defined as ideation in the past two weeks or a

lifetime attempt. For most sites, the PSS-3 was administered during the primary nursing assessment after the patient was placed in a treatment area. In the implementation phase of ED-SAFE 2, the local Lean Teams led efforts to examine screening maintenance and re-invigorate efforts to screen every ED patient for suicide risk as part of routine care. The local Lean Teams used several implementation strategies, including electronic health record integration (ranging from simple integration of the screening items only through programming of "hard stops" that mandated screening completion), ongoing training, data reporting and feedback, and spot checks with frontline staff.

2.2.3. The Safety Planning Intervention (SPI)

The SPI is a personalized approach that focuses on early identification of warning signs and execution of systematic steps to manage suicidal thoughts. The steps used by the suicidal individual include: [1] recognizing warning signs of an impending suicidal crisis, [2] employing internal coping/distracting strategies, [3] using social contacts and social settings as a means of distraction, [4] contacting family members or friends who may help to resolve the crisis, [5] contacting mental health professionals or agencies, and [6] reducing access to potentially lethal means.

For the ED-SAFE 2, a site trainer and clinicians were trained in-person by a Master SPI Trainer (Stanley or Brown). The in-person workshop focused on skills necessary to implement safety planning and utilized didactics, review of the manual, live and/or recorded modeling demonstration, and participant role-playing. Each site trainer was tasked with training other clinicians and helping to oversee implementation of the SPI. Regular site supervision calls with the Master SPI Trainer occurred after the initial site visit for the entire Implementation and Maintenance periods.

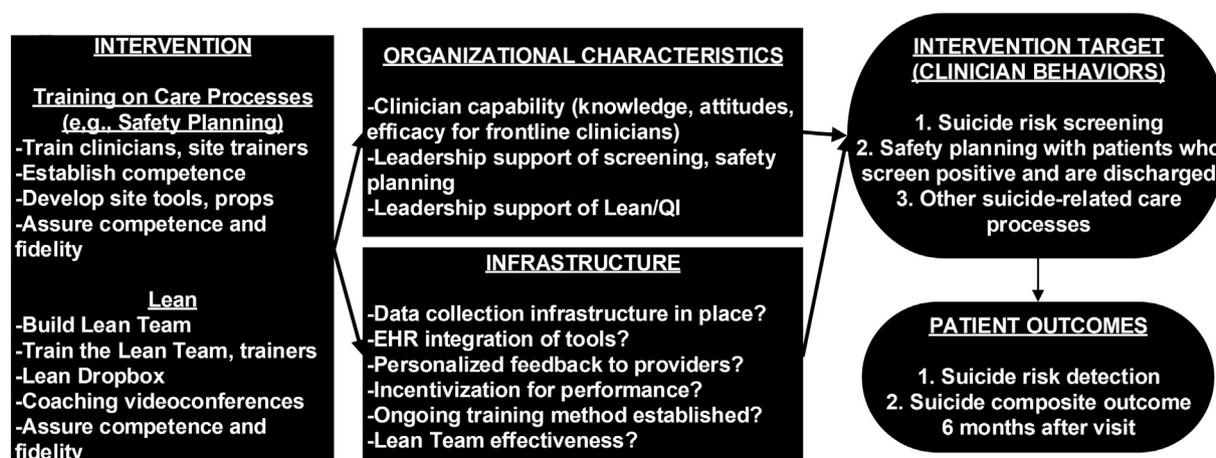
Competence in the SPI for the site trainer was determined through role playing of a standardized patient, a common method for establishing competence when actual observation or recordings with patients are impossible. The site trainers administered the SPI to a local partner, who played the role of a standardized patient. The role-play was audio-recorded and then independently rated by the site's Master SPI Trainer. All ratings were conducted using the Safety Plan Intervention Rating Scale (SPIRS), a standardized rating scale developed to ascertain competency in the intervention [25]. The SPIRS total score ranges from 0 to 20. All site trainers who obtained an SPIRS total score of at least 16 were identified as competent to provide training. Those not reaching this level were provided with feedback on their SPIRS ratings, re-trained by the Master SPI Trainers, and re-evaluated until the designated score was achieved. Newly hired staff were trained by the site trainer or remotely by Master SPI Trainers. Site trainers determined how to train and supervise on administration of the SPI.

2.3. Study sites

All eight of the original ED-SAFE 1 EDs participated, representing a diverse array of size, demographics, and geography [10].

2.4. Data collection and measurement

Fig. 1 depicts the hypothetical model predicting screening, risk detection, and clinician behaviors like safety planning, and patient outcomes. Unfortunately, there are no widely accepted, standardized measures for several constructs in the model, such as fidelity to Lean and clinician knowledge, attitudes, and efficacy related to the SPI. The Project Team took a practical approach to operationalizing these constructs. Where possible, recommended scale construction principles were used [26], including detailed construct definitions to guide item generation; vetting and revising items iteratively; piloting and revising items; and establishing relevant psychometrics, including internal consistency for survey scales.



Caption: The model describes how intervention components are hypothesized to influence care processes and patient outcomes

Fig. 1. Heuristic Model.

Caption: The model describes how intervention components are hypothesized to influence care processes and patient outcomes.

2.4.1. System outcomes

Documents submitted by the Lean Team and interviews with sites were used to establish the infrastructure support for suicide screening and safety planning. Specifically, sites were characterized based on the following: [1] a data collection infrastructure to support routine monitoring of performance, documented through the results boards, [2] support within the EHR for screening and safety planning, [3] personalized performance feedback to providers, [4] incentivized screening and safety planning, [5] ongoing training methods established, and [6] Lean Team effectiveness. The Lean A3 plans from each site were rated independently by Master Lean coaches (Pelletier, Johnson) using quality ratings that contained descriptive anchors aligning to “Needs Improvement, Satisfactory, or Excellent.” The Lean Team at each site also completed monthly self-reports. The first 5 elements were assessed using these data. For the sixth, team effectiveness, a self-administered survey of Lean Team members assessed team effectiveness, perceived team skill and goal agreement, and perceived organizational support. Questions from existing surveys were adapted [27], and the survey was carried out twice, at the end of Implementation and Maintenance phases.

2.4.2. Clinician measures

In addition to the Lean measures, clinician and organizational characteristics were assessed by surveying clinicians to establish: [1] clinician capability, measured by frontline staff knowledge, attitudes, and efficacy (3 subscales) around screening and safety planning; [2] departmental leadership support for screening and safety planning (1 subscale); and [3] departmental leadership support for CQI efforts, in general (1 subscale). At three time points – immediately before Implementation, midway through Implementation, and midway through Maintenance – all ED clinicians voluntarily completed a self-administered survey on paper or online. Responses for the five subscales were averaged within each clinician type (physician, nurse, mental health professional), yielding 15 organization characteristic measures at three time points, or 45 total for the study.

In addition, clinician behavior across a range of suicide-related care processes were measured, with a particular focus on those described below.

i. Suicide screening. For each month, 25 general ED patients ≥ 18 years old were randomly chosen to allow examination of universal screening rates of the general populace, resulting in 2400 total charts ($25 \times 12 \times 8 = 2400$) for each phase, 7200 charts total. A completed suicide screen was defined as any chart documentation of the presence OR absence of suicidal ideation AND presence or absence of a past suicide attempt (0 = None documented, 1 = Partial (+/- ideation OR

attempt noted), 2 = Complete (+/- ideation AND attempt noted). The proportion of general ED patients with a complete screen within each study phase comprised the primary clinician-level outcome for analyses pertaining to screening. Sensitivity analyses will expand the definition to partial OR complete screens.

ii. Safety planning. To assess safety planning, 25 additional patients per month with a positive suicide risk screen were randomly chosen and reviewed. The review identified those who were discharged home and, therefore, eligible to receive an SPI. The original ED-SAFE 1 found that 25% of a screen-positive sample was discharged. Consequently, while 2400 suicidal patient charts were reviewed, only about 600 discharged individuals per phase (1800 total) were available for examination of safety planning specifically.

A safety plan is defined as a written document provided to the patient that specifies steps that he/she should follow when feeling suicidal. It was coded as 0 = None, no documentation of a safety plan, 1 = Partial, documentation of safety planning in chart text, but no actual copy in the record, or 2 = Complete, copy of the safety plan in the record. When present, the safety plans were quality coded. The proportion of discharged screen-positive patients with a complete safety plan within each study phase will comprise the primary clinician-level outcome for analyses pertaining to safety planning. Sensitivity analyses will expand the definition to partial OR complete safety plans.

2.4.3. Patient measures

Patient outcomes were gathered through multiple methods: [1] random chart reviews, [2] healthcare utilization reviews, [3] National Death Index (NDI) queries, [4] clinician surveys from three time-points, [5] Lean Team monthly surveys and documents, and [6] qualitative interviews with key informants conducted across the entire study period.

i. Suicide risk detection. Using the parent ED-SAFE 1 criteria, a positive screen is defined as any individual who either endorsed active suicidal ideation OR reported a lifetime suicide attempt. The proportion of general ED patients with a positive screen within each study phase will comprise the primary patient-level outcome for analyses pertaining to risk detection.

ii. Suicide composite outcome. Each subject not only had the randomly selected index visit reviewed but was also followed through all electronic health records associated with the site's healthcare system for six months to document all suicide-related outcomes, including ED visits and hospitalizations related to suicidal ideation, suicidal attempts, or completed suicide. The records were reviewed and abstracted at the site.

The NDI was queried to determine deaths in the six months after the

index visit. If someone had died, the probability of suicide from the death record was abstracted. These two sources (medical record review, NDI) will be combined to construct the suicide composite outcome, which will be defined as an ED visit or hospitalization due to suicidal ideation/behavior OR death by suicide in the six months after the index visit. Standard definitions established by the CDC [19] and the original ED-SAFE 1 will be used. The proportion of patients who screened positive for suicide risk but were discharged and ended up with a suicide composite outcome over the six months after the ED visit will be the primary patient outcome for evaluating impact on suicidal behavior.

2.5. Analysis plan

As of submission of this paper, the study has been completed, but the data have not yet been analyzed.

2.5.1. Qualitative analytic strategy (implementation facilitators)

Qualitative data will be used to provide additional depth surrounding the implementation process, barriers and solutions faced by the sites, similarities and differences across sites, and lessons learned. All key informant interviews will be digitally recorded for verbatim transcription. Analysis of qualitative data will be conducted using ATLAS.ti, a qualitative data analysis software program allowing fluid “interaction” of data across types and sources. Initially, a top-level overview will be created based on the interview guide and the conceptual framework. The codebook will be elaborated upon based on emergent themes, and adjusted as interviews proceed. Interviews will be compared within each site, within each sampling sub-frame, and across each group. Analyses across transcripts will identify the most salient components in the model depicted in Fig. 1. Theme summaries will be compared against the quantitative data and discrepancies will be reconciled through additional interviews.

2.5.2. Quantitative analytic strategy (suicide-related care and outcomes)

The study will examine suicide-related care processes and how they relate to suicidal behavior after the index ED visit. The approach to evaluate sustainability of universal suicide risk screening in adult ED patients at Phase 1 (Baseline) and again after Phase 3 (Maintenance) will be used as an example. Clinician screening behavior and the patient suicide risk detection (positive screens) among all ED patients will be plotted over the three phases. Factors affecting the maintenance of clinician screening between the last phase of the original study and the Baseline (follow) period will be examined using a generalized estimating equation (GEE) model. The ED site will be used as a clustering factor and the model will take into account: ED-level characteristics, including a time indicator for the original data and the new Baseline data; the Step for site initiation (i.e., Step 1, 2, 3, 4); organizational characteristics and infrastructure indicators listed in Fig. 1, collected during the original parent study; and patient characteristics, e.g., gender, age. This model will identify factors predictive of change in screening rate from the end of the original parent study to the Baseline phase and, by extrapolation, factors predictive of maintenance (i.e., no change in screening rate). A second model will extend the first model and investigate factors predictive of suicide risk detection (i.e., a positive screen for suicide risk, the patient-level outcome), including screening as a mediator for risk detection.

The second analytic approach will use the GEE logistic models in an interrupted time series paradigm [28] to investigate the slope of the screening rate during the final phase of the original ED-SAFE 1 study, the slope of screening rate during ED-SAFE 2 Baseline, and the change in rate associated with the completion of the original study and phasing out of funding support. The modeling should allow identification of site characteristics, mechanistic factors, and patient subgroups associated with a significantly lower (or higher) probability of sustained screening and risk detection. Statistical approaches to sustainability of screening and other care processes like safety planning are very similar, and thus

not discussed in detail. For example, as above, safety planning (clinician behavior) and suicide-related composite outcome will be plotted across all three phases, and GEE will be used to evaluate changes and to model the factors that relate to implementation and maintenance success. All analyses will be conducted using SAS (SAS 9.4, Institute, Cary, NC).

i. Sample size consideration. PASS 11 (Kaysville, Utah) was used to estimate sample size. The comparison used to estimate sample size was the safety planning rate for Baseline vs. Implementation. This was chosen because it is a less common outcome than screening; powering for safety planning should provide ample power to detect changes in the more common behavior of screening. While 2400 screen-positive charts were reviewed per phase, only 25% (600/phase) were expected to be discharged and, therefore, eligible for a safety plan. Because no closed form sample size equations exist for a hierarchical GEE logistic model, the initial simplifying assumption collapses the levels of the hierarchy.

Although it is difficult to know exactly what the rate of safety planning will be during the Baseline phase, original ED-SAFE 1 data based on chart reviews and interviews with clinicians and patients suggests it will be less than 25%. However, it is important to use a conservative estimate for Baseline to protect against under-powering. Consequently, a value of 50% safety planning rate was assumed for Baseline, because 50% is the worst-case estimate that yields the largest standard deviation. In addition, the absolute increase in safety planning was fixed at a conservative level of 15% increase, making the expected rate during Implementation 65%. Since the a priori performance target is 90%, this effect size is also very conservative. A logistic regression of the binary response variable (safety plan: yes/no) on the binary predictor variable (treatment: Baseline/Implementation) with a sample size of 1200 observations of discharged patients (600 Baseline/600 Implementation) achieves 90% power at a 0.05 significance level to detect a change in the probability of receiving a safety plan from the Baseline value of 50% to 65%. This corresponds to an odds ratio of 1.86.

2.6. Human participant considerations

Because this was an implementation study, patients did not sign informed consent for chart reviews but did complete informed consent procedures for fidelity interviews. Clinicians who completed the surveys were provided a description of the survey and they could opt out of completion.

3. Intervention fidelity

3.1. Screening fidelity

As in the original ED-SAFE 1, to measure the fidelity of suicide screening implementation (e.g., determine whether nursing staff are actually asking the suicide screening questions vs. simply “intuiting” them and documenting that screening was done), research assistants (RAs) completed a brief interview with a random sample of patients whose medical record indicated a negative screen (i.e. the chart documents an absence of suicidal ideation or behavior). At each site, this comprised 30 patients during the first and last months of Implementation ($n = 60$) and 30 patients during the first and last months of Maintenance ($n = 60$). Fidelity was represented by the proportion of patients interviewed who reported having been asked the suicide screening questions. In those patients who reported they were not asked the questions, the RA administered the suicide risk screener to help establish the potential clinical impact of poor fidelity (false negatives). Published protocols for informing clinicians of new positive screens ensured participant safety [29–30]. In the original ED-SAFE study, fidelity ranged from 65 to 95%, improving over time [7].

3.2. Safety planning intervention fidelity

All clinical staff were trained to post the safety plan to the medical record to ensure it could be identified by the RAs whenever it was done. During the random chart reviews, the safety plans were coded by RAs using a strategy developed by Master SPI Trainers that rated the six steps as satisfactory [31]. While not a requirement, many sites integrated the safety planning template into their EHRs, which further facilitated fidelity measurement.

3.3. Lean fidelity

Fidelity to Lean was measured by Lean Team documents that support: [1] the site having updated their A3 to document completion of two additional improvement cycles, [2] representation from at least one ED leader and frontline employee at all Lean Team meetings, and [3] updating of the results board after every improvement cycle. These components are emblematic of Lean tools and constituted the core of the Lean intervention.

4. Conclusion

The results will build on the remarkable success of the original ED-SAFE 1 and will have a sustained public health impact through promoting better suicide risk detection and suicide prevention. It will promote rigorous models of evaluating QI efforts and study mechanisms of action firmly rooted in implementation science theory. It will advance research methods, including analyses of stepped wedge and interrupted time series trial designs, which are broadly relevant to pragmatic clinical trials. These methods will allow us to explore the mechanisms by which the implementation effort was successful or why it failed. Finally, because other settings face similar challenges in studying routine suicide screening and prevention [32], the results will have implications beyond the walls of the ED, and can be adapted to primary care, inpatient, and specialty settings.

Funding

The work was supported by the National Institute of Mental Health [Award Number R01MH106726]. 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 ED-SAFE 2 Investigator Team:

Ivan W. Miller, PhD; Departments of Psychiatry and Human Behavior, Brown University, Providence, RI, USA, Ivan_Miller_III@brown.edu

Sarah A. Arias, PhD; Departments of Psychiatry and Human Behavior, Brown University, Providence, RI, USA, sarah.arias@brown.edu

Carlos A. Camargo, Jr., MD, DrPH; Department of Emergency Medicine, Massachusetts General Hospital, Boston, MA, USA, ccamargo@partners.org

Maura Kennedy, MD; Department of Emergency Medicine, Massachusetts General Hospital, Boston, MA, USA, mkennedy@bidmc.harvard.edu

Laura Macnow, MD; Department of Emergency Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA, lmacnow@bid.harvard.edu

Katherine Boyle, MD; Department of Emergency Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA, klboyle@bidmc.harvard.edu

Frank LoVecchio, DO, MPH, FACEP; Departments of Emergency Medicine and Medical Toxicology, Banner University Medical Center, Phoenix, AZ, USA, Frank.LoVecchio@bannerhealth.com

Lisa Uebelacker, PhD; Departments of Psychiatry and Human Behavior, Brown University, Providence, RI, USA, Lisa_Uebelacker@brown.edu

Jeffrey Caterino, MD, MPH; Department of Emergency Medicine, The Ohio State University Wexner Medical Center, Columbus, OH, USA Jeffrey.Caterino@osumc.edu

Carly Eastin, MD; Department of Emergency & Trauma, University of Arkansas for Medical Sciences, Little Rock, AR, USA, CDEastin@uams.edu

Marian E. Betz, MD, MPH; Department of Emergency Medicine, University of Colorado Denver – Anschutz Medical Campus, Denver, CO, USA, marian.betz@cuanschutz.edu

Wesley Zeger, DO; Department of Emergency Medicine, University of Nebraska Medical Center, Omaha, NE, USA, wzeger@unmc.edu

Ameer F. Ibrahim, MD; Department of Emergency Medicine, University of Massachusetts Medical School, Worcester, MA, USA, ameer.ibrahim@umassmemorial.org

References

- [1] M.E. Betz, R. Schwartz, E.D. Boudreaux, Unexpected suicidality in an older individual in an emergency department, *J. Am. Geriatr. Soc.* 61 (6) (Jun 2013) 1044–1045, <https://doi.org/10.1111/jgs.12290>.
- [2] B. Stanley, G.K. Brown, Safety planning intervention: a brief intervention to mitigate suicide risk, *Cogn. Behav. Pract.* 19 (2012) 256–264, <https://doi.org/10.1016/j.cbpra.2011.01.001>.
- [3] J.A. Rings, P.A. Alexander, V.N. Silvers, P.M. Gutierrez, Adapting the safety planning intervention for use in a Veterans' psychiatric inpatient group setting, *J. Ment. Health Couns.* 34 (2012) 95–109, <https://doi.org/10.17744/mehc.34.2.a77036631424nmq7>.
- [4] H. Meyer, Life in the 'lean' lane: performance improvement at Denver health, *Health Aff (Millwood)*. 29 (11) (Nov 2010) 2054–2060, <https://doi.org/10.1377/hlthaff.2010.0810>.
- [5] J. Toussaint, R. Gerard, *On the Mend: Revolutionizing Healthcare to Save Lives and Transform the Industry*, Lean Enterprise Institute, Inc., 2010.
- [6] R.J. Holden, Lean thinking in emergency departments: a critical review, *Ann. Emerg. Med.* 57 (3) (Mar 2011) 265–278, <https://doi.org/10.1016/j.annemergmed.2010.08.001>.
- [7] E.D. Boudreaux, C.A. Camargo Jr., S.A. Arias, et al., Improving suicide risk screening and detection in the emergency department, *Am. J. Prev. Med.* 50 (4) (Apr 2016) 445–453, <https://doi.org/10.1016/j.amepre.2015.09.029>.
- [8] B.K. Ahmedani, G.E. Simon, C. Stewart, et al., Health care contacts in the year before suicide death, *J. Gen. Intern. Med.* 29 (6) (Jun 2014) 870–877, <https://doi.org/10.1007/s11606-014-2767-3>.
- [9] E.D. Boudreaux, I. Miller, A.B. Goldstein, et al., The emergency department safety assessment and follow-up evaluation (ED-SAFE): method and design considerations, *Contemp Clin Trials*. 36 (1) (Sep 2013) 14–24, <https://doi.org/10.1016/j.cct.2013.05.008>.
- [10] I.W. Miller, C.A. Camargo Jr., S.A. Arias, et al., Suicide prevention in an emergency department population: the ED-SAFE study, *JAMA Psychiatry*. 74 (6) (2017) 563–570, <https://doi.org/10.1001/jamapsychiatry.2017.1776>.
- [11] Joint Commission, Detecting and treating suicide ideation in all settings, *Sentinel Event Alert* 56 (2016) 1–7.
- [12] B. Stanley, G.K. Brown, L.A. Brenner, et al., Comparison of the safety planning intervention with follow-up vs usual care of suicidal patients treated in the emergency department, *JAMA Psychiatry*. 75 (9) (2018) 894–900, <https://doi.org/10.1001/jamapsychiatry.2018.1776>.
- [13] D.A. Chambers, R.E. Glasgow, K.C. Stange, The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change, *Implement. Sci.* 8 (117) (2013) 1–11, <https://doi.org/10.1186/1748-5908-8-117>.
- [14] T. Greenhalgh, G. Robert, F. Macfarlane, P. Bate, O. Kyriakidou, Diffusion of innovations in service organizations: systematic review and recommendations, *Milbank Q.* 82 (4) (2004) 581–629, <https://doi.org/10.1111/j.0887-378X.2004.00325.x>.
- [15] G.A. Aarons, M. Hurlburt, S.M. Horwitz, Advancing a conceptual model of evidence-based practice implementation in public service sectors, *Admin. Pol. Ment. Health* 38 (1) (Jan 2011) 4–23, <https://doi.org/10.1007/s10488-010-0327-7>.
- [16] Stirman Wiltsey, S. Kimberly, J. Cook, N. Calloway, A. Castro, F. Charns, M., The sustainability of new programs and innovations: a review of the empirical literature and recommendations for future research, *Implement. Sci.* 7 (2012) 17, <https://doi.org/10.1186/1748-5908-7-17>.
- [17] K. Hemming, T.P. Haines, P.J. Chilton, A.J. Girling, The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting, *BMJ*. 350 (2015) h391, <https://doi.org/10.1136/bmj.h391>.
- [18] H.C. Kaplan, L.P. Provost, C.M. Froehle, P.A. Margolis, The model for understanding success in quality (MUSIQ): building a theory of context in healthcare quality

- improvement, *BMJ Qual. Saf.* 21 (1) (Jan 2012) 13–20, <https://doi.org/10.1136/bmjqs-2011-000010>.
- [19] H.C. Kaplan, C.M. Froehle, A. Cassidy, L.P. Provost, P.A. Margolis, An exploratory analysis of the model for understanding success in quality, *Health Care Manage Rev Oct-Dec* 38 (4) (2013) 325–338, <https://doi.org/10.1097/HMR.0b013e3182689772>.
- [20] A.C. Feldstein, R.E. Glasgow, A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice, *Jt Comm J Qual Patient Saf.* 34 (4) (Apr 2008) 228–243, [https://doi.org/10.1016/S1553-7250\(08\)34030-6](https://doi.org/10.1016/S1553-7250(08)34030-6).
- [21] L.J. Damschroder, D.C. Aron, R.E. Keith, S.R. Kirsh, J.A. Lowery Alexander, J.C., Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science, *Implement. Sci.* 4 (2009) 50, <https://doi.org/10.1186/1748-5908-4-50>.
- [22] M.A. Ilgen, M.A. Walton, R.M. Cunningham, et al., Recent suicidal ideation among patients in an inner city emergency department, *Suicide Life Threat. Behav.* 39 (5) (Oct 2009) 508–517, <https://doi.org/10.1521/suli.2009.39.5.508>.
- [23] G.L. Larkin, A.L. Beautrais, Emergency departments are underutilized sites for suicide prevention, *Crisis.* 31 (1) (2010) 1–6, <https://doi.org/10.1027/0227-5910/a000001>.
- [24] E.D. Boudreaux, M.L. Jaques, K.M. Brady, A. Matson, M.H. Allen, The patient safety screener: validation of a brief suicide risk screener for emergency department settings, *Arch Suicide Res.* 19 (2) (2015) 151–160, <https://doi.org/10.1080/13811118.2015.1034604>.
- [25] Brown, G. Stanley, B. Safety plan intervention rating scale: University of Pennsylvania and Columbia University; unpublished manual.
- [26] E.J. Pedhazur, L. Pedhazur-Schmelkin, *Measurement, Design, and Analysis: An Integrated Approach*, Psychology Press, Taylor & Francis Group, New York, NY, 1991.
- [27] S.M. Shortell, J.A. Marsteller, M. Lin, et al., The role of perceived team effectiveness in improving chronic illness care, *Med. Care* 42 (11) (2004) 1040–1048 Nov <https://www.jstor.org/stable/4640854>.
- [28] S.L. Zeger, R. Irizarry, R.D. Peng, On time series analysis of public health and biomedical data, *Annu. Rev. Public Health* 27 (2006) 57–59, <https://doi.org/10.1146/annurev.publhealth.26.021304.144517>.
- [29] M.H. Allen, B.W. Abar, M. McCormick, et al., Screening for suicidal ideation and attempts among emergency department medical patients: instrument and results from the psychiatric emergency research collaboration, *Suicide Life Threat. Behav.* 43 (2013) 313–323, <https://doi.org/10.1111/sltb.12018>.
- [30] E.D. Boudreaux, M.H. Allen, C. Claassen, et al., The psychiatric emergency research collaboration-01: methods and results, *Gen. Hosp. Psychiatry* 31 (6) (2009) 515–522, <https://doi.org/10.1016/j.genhosppsych.2009.04.009>.
- [31] Brown, G. Stanley, B. Coding manual for rating of safety plans in medical records: University of Pennsylvania and Columbia University; unpublished manual.
- [32] E. O'Connor, B.N. Gaynes, B.U. Burda, C. Soh, E.P. Whitlock, Screening for and treatment of suicide risk relevant to primary care: a systematic review for the U.S. Preventive Services Task Force, *Ann. Intern. Med.* 158 (10) (May 21 2013) 741–754, <https://doi.org/10.7326/0003-4819-158-10-201305210-00642>.