

Suicide Behavior Results From the U.S. Army's Suicide Prevention Leadership Tool Study: The Behavioral Health Readiness and Suicide Risk Reduction Review (R4)

LTC Justin M. Curley, MD, USA; Farifteh F. Duffy, PhD; Paul Y. Kim, MA; Kristina M. Clarke-Walper, MPH; Lyndon A. Riviere, PhD; Joshua E. Wilk, PhD

ABSTRACT

Introduction:

The U.S. Army developed a new tool called the Behavioral Health Readiness and Suicide Risk Reduction Review (R4) for suicide prevention. A 12-month evaluation study with the primary objective of testing the hypothesis (H1) that Army units receiving R4 would demonstrate improved outcomes in suicidal-behavior measures following the intervention, relative to control, was then conducted. The results of analyses to answer H1 are herein presented.

Materials and Methods:

The R4 intervention (R4-tools/instructions/orientation) evaluation study, Institutional Review Board approved and conducted in May 2019-June 2020, drew samples from two U.S. Army divisions and employed a repeated measurement in pre-/post-quasi-experimental design, including a nonequivalent, but comparable, business-as-usual control. Intervention effectiveness was evaluated using self-report responses to suicide-related measures (Suicide Behaviors Questionnaire—Revised/total-suicide behaviors/ideations/plans/attempts/non-suicidal self-injuries) at 6-/12-month intervals. Analyses examined baseline to follow-up linked and cross-sectional cohorts, incidence/prevalence, and intervention higher-/lower-use R4 subanalyses.

Results:

Both divisions demonstrated favorable in-study reductions in total-suicide burden, with relatively equivalent trends for total-suicide behaviors, total-suicide risk (Suicide Behaviors Questionnaire—Revised), suicidal ideations, and non-suicidal self-injuries. Although both demonstrated reductions in suicide plans, the control showed a more robust trend. Neither division demonstrated a significant reduction in suicide attempts, but subgroup analyses showed a significant reduction in pre-coronavirus disease 2019-attempt incidence among those with higher-use R4 relative to control.

Conclusions:

There is no evidence of harm associated with the R4 intervention. R4 effectiveness as a function of R4 itself requires confirmatory study. R4 is judged an improvement (no evidence of harm + weak evidence of effectiveness) over the status quo (no safety data or effectiveness studies) with regard to tool-based decision-making support for suicide prevention in the U.S. Army.

INTRODUCTION

Over the course of 2018, the U.S. Army developed a new tool called the Behavioral Health Readiness and Suicide Risk Reduction Review (R4) to confront the challenge of suicide in the force.¹ R4 development integrated leadership feedback with evidence-based predictors of suicide risk, taking into account both leader practices and empirical findings.

Department of Military Psychiatry, Walter Reed Army Institute of Research, Silver Spring, MD 20910, USA

The material has been reviewed by the Walter Reed Army Institute of Research, in addition to the offices associated with the listed authors. There is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army, Department of Defense, or the U.S. Government. The investigators have adhered to the policies for the protection of human subjects as prescribed in Army Regulation 70–25.

doi:<https://doi.org/10.1093/milmed/usac169>

Published by Oxford University Press on behalf of the Association of Military Surgeons of the United States 2022. This work is written by (a) US Government employee(s) and is in the public domain in the US.

The resultant R4 tools and processes equipped Army leaders with empirically derived knowledge and workflows designed to improve suicidal-behavior outcomes.

At the individual level, R4 techniques and procedures emphasized the paired identification/awareness of risk factors with leader responses that promoted engagement and communication between the leader and led. At the unit level, R4 design informed and mobilized timely management and resourcing of at-risk soldiers to reduce risk at each echelon of the chain of command. Together, the R4 tools and processes provided U.S. Army leaders with practical methods for earlier identification and optimization of the health and welfare of soldiers who may be at risk of suicide, while simultaneously enhancing and synchronizing the processes necessary to support and care for soldiers through existing Army systems.

Following R4's development, a 12-month evaluation study of the R4 intervention with two active duty Army divisions was then conducted.² Despite the use of many previous tools and programs for suicide prevention, this was the first time the Army was able to empirically test the effectiveness of

tool-supported decision-making among Army units in a rigorous fashion. The primary objective of the R4 study was to test the hypothesis that Army units within the division receiving the R4 intervention would demonstrate improved outcomes in suicidal-behavior measures following intervention, relative to a comparable control division. This article will provide the results of the analyses conducted to answer the main hypothesis of the R4 study.

METHODS

The R4 study was approved by the Walter Reed Army Institute of Research's Institutional Review Board and conducted by the Walter Reed Army Institute of Research. A detailed description of R4 study methodology has been previously published.² The R4 intervention and evaluation was implemented in May 2019 and concluded in June 2020. The R4 intervention consisted of R4 tools, accompanying instructions, and an orientation. These were packaged in a tiered fashion based on leadership echelon (platoon upward to division).

The tools and accompanying instructions—one for platoon-level (ranks E5-E7 and lieutenants) and one for company commander/first sergeant end-users—were provided to the corresponding leaders within the context of a 30-50-minute orientation session.^{1,2} The standardized intervention material included a brief overview of the challenge of suicide in the U.S. Army, the importance of suicide prevention leader decision-making in the force, R4 ingredients for making such decisions (i.e., identification of risk factors paired with face-to-face conversations by leaders with soldiers to assess risk—also called “engaged leadership”), an introduction to operating whichever of the two instruments was issued, a demonstration on how to use R4 based on the position in the organization (i.e., management, resourcing, and readying of soldiers who might be at risk of suicide or negative behavioral health outcomes), vertical/horizontal chain-of-command synchronization of effort, and two individual example scenarios whereby leaders could use R4 to enhance their suicide prevention decision-making.

R4 implementation involved (1) a one-time “direct” training of the intervention site's end-user leaders (Sergeant [E5] and above) by R4 staff in May 2019 and (2) “indirect” ongoing diffusion via leaders' sharing R4 tools, instructions, and processes with R4 naïve (new or unavailable for direct training) intervention site end-users throughout the 12-month evaluation period. The evaluation of the R4 intervention employed a repeated measurement in pre-/post-quasi-experimental design, including a nonequivalent, but comparable, business-as-usual control group. Samples were drawn from two geographically separated U.S. Army divisions in the continental United States, each composed of four comparable combat brigades. Soldiers in both intervention and control groups completed anonymous survey instruments to assess a range of psychological and physical health factors.²

The primary focus of the inquiry herein was the longitudinal assessment of intervention and control divisions necessary to answer the main hypothesis. To that end, primary analyses examined the soldier cohort with a baseline assessment (T1) who could be linked within their division of assignment at 6-month (T2) and/or 12-month follow-up (T3). This ensured that only participants subject to the potential effects of the intervention and control conditions across time points were included in primary analyses. The secondary cross-sectional analyses examined the dynamics of the larger population trends at each time point.

Sample

R4 study soldiers participated in accordance with an Army operations order, and each individual was provided the option to consent to the use of their data for research purposes after being briefed on the nature of the study. Participants were linked across time using a code derived from responses on personal (but non-identifying) questions such as state where they graduated high school and day of the month they were born.³ All surveys were anonymous and excluded all personally identifying information. Individual responses were kept confidential. No survey results were briefed to commands during the 12-month study time frame, and only aggregated results were briefed to commands following study completion.

At T1, 6,747 soldiers in the intervention division and 7,269 in the control division participated in data collection, with 4,705 (69.7%) and 5,275 (72.6%), respectively, providing consent. At T2, 8,013 and 8,541 participated, with 6,914 (86.3%) and 7,441 (87.1%), respectively, consenting. At T3, 9,175 and 7,177 participated, with 7,782 (84.8%) and 5,907 (82.3%), respectively, consenting.

At T2, data for 4,622 soldiers who remained within their respective divisions were linked with T1. Among those, 690 soldiers did not consent and were removed from analyses, leaving 3,932 as the focus of T1-to-T2 analyses. By T3, consented soldiers with the T1-T3 linkage was 2,655, which served as the primary focus of T1-to-T3 analyses. A cohort of 5,971 minus 860 non-consenters (5,111 total; 2,359 intervention; 2,752 control) could be linked across T1-T2, T1-T3, and T1-T2-T3. This group was analyzed to determine combined new events in the T1-T3 study time frame.

Supplementary Table S1 provides demographic and military characteristics of consented soldiers in each division, both cross-sectional and linked T1-to-T2 and T1-to-T3. The equivalence of the intervention division was assessed based on the comparability of its demographic and military characteristics with (1) the R4 control division and (2) the overall U.S. Army's organizational structure.⁴

Measures

Suicidal behaviors and suicide risk

The effectiveness of the R4 intervention was evaluated using self-report responses to a battery of suicide-related measures

at the intervals mentioned earlier. Suicidal behaviors consisted of ideations, plans, and attempts. Non-suicidal self-injury (NSSI) was also assessed but not included in suicidal-behavior totals. The occurrence of each behavior was assessed via Yes/No answers during the pre-T1, T1-T2, T1-T3 time points, and lifetime. The change in soldier outcomes over time (1) compared the change from baseline to follow-up among soldiers in the intervention group (within) and (2) contrasted the change over time for soldiers in intervention and control groups (between). These analyses were conducted using linked data, at the division-level, and also by R4-use status (as mentioned later).

The 12-month prevalence of suicidal behavior(s) at T1 and T3 and 6-month prevalence at T2 were computed for each time point. The 6-month incidence of suicidal behavior(s) at T2 was computed among soldiers with no reported lifetime suicidal behavior(s) at T1. Among soldiers with no reported prior suicidal behavior at T1, any report of suicidal behavior(s) since R4 implementation (T2 and/or T3) was tallied to determine any occurrence of new suicidal behavior(s) during the 12-month study period.

Data on a modified version of the Suicide Behaviors Questionnaire—Revised (SBQ-R) were also collected to assess individuals at risk for suicidal behaviors. The SBQ-R consisted of four items to assess different dimensions for determining suicide risk: (1) lifetime suicidal ideation, plan, and attempts; (2) frequency of suicidal ideation in the previous year; (3) suicidal threats; and (4) the likelihood of future suicide attempts. A score of ≥ 7 met the criteria for individual risk.⁵

Measures of R4-intervention reach and use

Since the implementation of the R4 intervention employed direct and indirect training approaches, R4 dissemination throughout the intervention division, defined herein as “reach,” was measured based on the total number of soldiers who completed the R4 evaluation at the intervention site at each time point. The proportion “reached” is based on the total number reached at each time point divided by 13,100 (theoretical estimate of total soldiers assigned within the sample’s division-level construct at any given time point).^{2,4}

The use/nonuse of R4 was measured based on unit leaders’ reports at T2 and T3 on (1) personally utilizing R4 tools/processes during the prior week, month, quarter, or since May 2019 for “none” to “1-2” or greater number of soldiers in their respective units and (2) affirming the unit’s utilization of R4 processes to identify, manage, and resource soldiers who experienced difficulties or may be at risk for suicide.

An additional measure of “R4-use” was created for the T1-T2 and T1-T2-T3 periods to account for the extent (quantity and quality) of leader training received as well as the use of R4 tools/processes in intervention battalions (BNs). A detailed description of R4 train-use variable development can be found in Supplementary Materials and Methods. In summary, based on leaders’ R4 training and use reports, each intervention

BN was assigned to “high-use R4” and “low-use R4” by T2 and “high-use R4,” “intermediate-use R4,” and “low-use R4” by T3.

Suicidal behaviors post-coronavirus disease 2019

The implementation of the T3 data collection overlapped with the coronavirus disease 2019 (COVID-19) pandemic for approximately the final quarter of the 12-month study time frame. A series of questions inquired about the change in frequency of soldier experiences with any potential suicidal thoughts or attempts since the COVID-19 pandemic (i.e., February 2020 to study completion), compared to before. Response options included “not applicable,” “decreased,” “stayed about the same,” and “increased.”

Analytic Approach

All analyses were performed using SPSS v24. Basic frequencies and descriptive analyses were used to document the reach and use of the R4 tool/processes within the intervention group and to inform decision rules for subsequent analyses.

Main analyses examined the cohort of soldiers with baseline to follow-up linkage. In this manner, the full potential effects of the intervention and control conditions could be studied among soldiers with linked data across time points. Pairwise deletion was used for these analyses.

Pearson chi-squared and Fisher’s exact tests were used to examine bivariate relationships and a comparison of proportions for independent samples. The two-sample McNemar and Cochran’s Q tests were conducted for a comparison of pre-/post-responses for multiple group design, accounting for correlated proportions within each group.

Given anticipated sizable losses to follow-up due to availability and attrition² among soldiers with linked data, secondary analyses were conducted using cross-sections of unbalanced panel data for the three time points. Each cross-section included all soldiers present in all or some data collection time points, linked or not-linked, or new to data collection at T2 or T3. For preliminary analyses, Pearson chi-squared/Fisher’s exact tests were utilized for between-group comparison of proportions among the intervention and control groups at each time point. Additionally, Z-score tests for overlapping samples were conducted to compare within-group differences in proportions from T1 to T3 and separately for the intervention and control groups.⁶

RESULTS

Intervention Reach and Use at T1, T2, and T3

At T1, out of an estimated 13,100 potentially assigned soldiers, a total of 6,747 (51.5%) in the intervention division participated in data collection. Of those, 2,690 were unit leaders and the R4 training focus. At T2, 8,013 soldiers participated (61.2% of assigned), of which 3,506 were leaders. A total of 2,141 (61.1%) leaders reported having received R4, and 1,040 (29.7%) had used the R4 processes for one or more

soldiers in their units. At T3, of 9,175 participants (70.0% of assigned) 3,735 were leaders. A total of 2,566 (68.7%) leaders reported having received R4, and 1,408 (37.7%) had used R4 with one or more soldiers.

Demographic, Military, and Baseline Suicidal Behavior Characteristics: Comparing Intervention and Control Divisions

Demographic and military characteristics for T1, T2, and T3 cross-sectional, T1-to-T2 linked, and T1-to-T3 linked samples of soldiers who consented to research are presented in [Supplementary Table S1](#). Age distribution was comparable across divisions in all five samples. Race/ethnicity distribution was comparable across both linked samples. A higher proportion of soldiers in most intervention division samples reported being married, assigned to brigade combat teams, and had previously deployed for combat. A higher proportion of soldiers in most control samples were assigned to a sustainment brigade. Finally, at the cross-sectional T1 assessment, the intervention division began the study with a significantly higher number of reported suicide attempts relative to the control ($P < 0.01$, [Table II](#)).

Suicidal Behaviors Across Time

Any suicidal behavior

Based on analyses of T1-to-T3 linked data ([Table I](#)), smaller proportions of soldiers in both the intervention and control divisions reported suicidal behaviors (total ideations, plans, and attempts) by T3 when compared to T1: net decreases of 2.2% ($P < 0.05$) and 3.7% ($P < 0.001$), respectively. Neither the magnitude of this baseline-to-follow-up change across divisions nor the difference in proportions across divisions at T1 or at T3 was statistically significant.

Cross-sectional within-division results also demonstrated smaller proportions of soldiers in both the intervention and control divisions reporting suicidal behaviors by T3 when compared to T1: net decreases of 3.7% ($P < 0.001$) and 5.3% ($P < 0.001$), respectively ([Table II](#)). The smaller proportion of soldiers in the control division cross-sectional sample who reported suicidal behaviors in comparison to the intervention division was significant at T3 ($P < 0.001$).

Six-month incidence (i.e., T2 reports of new suicidal behaviors among soldiers without lifetime suicidal behaviors at T1) and the proportion of new suicidal behaviors (i.e., T2 and/or T3 reports of new suicidal behavior among soldiers without lifetime suicidal behaviors at T1) during the 12-month study period are presented in [Table III](#).

Suicide Behaviors Questionnaire—Revised

Based on analyses of T1-to-T3 linked data ([Table I](#)), smaller proportions of soldiers in both the intervention and control divisions demonstrated an SBQ-R score of ≥ 7 (SBQ-R7) during the 12-month time frame: net decreases of 2.5% ($P < 0.01$) and 2.8% ($P < 0.01$), respectively. Neither the magnitude of this baseline to follow-up change across divisions nor the

TABLE I. Two-Sample McNemar Test for 12-Month Suicidal Behaviors at T1 and T3, by Intervention Status (Linked Data)

Assignment	Suicidal behaviors																				
	Any suicidal behaviors Past 12 months				Any suicidal ideations Past 12 months				Any suicidal plans Past 12 months				Any suicide attempts Past 12 months				Any non-suicidal self-injury Past 12 months				
	T1	T3	%	n	T1	T3	%	n	T1	T3	%	n	T1	T3	%	n	T1	T3	%	n	
Intervention (N = 1,305)	13.6 ^a	11.4 ^a	178	149	13.0 ^a	10.9 ^a	170	142	5.3	69	4.8 ^{c*}	63	2.0	26	1.5	20	2.9 ^{a*}	38	1.8 ^{a*}	24	
Control (N = 1,350)	12.9 ^{b***}	9.2 ^{b***}	174	124	12.1 ^{b***}	8.8 ^{b***}	164	119	4.8 ^{b*}	65	3.1 ^{b*c*}	42	1.3	17	0.7	10	2.7 ^{b**}	37	1.4 ^{b**}	19	
	SBQ-R ≥ 7																				
Intervention (N = 1,305)	13.5 ^{a**}	11.0 ^{a**}	176	143																	
Control (N = 1,350)	11.5 ^{b**}	8.7 ^{b**}	155	117																	

Abbreviation: SBQ-R, Suicide Behaviors Questionnaire—Revised.

^aWithin intervention group McNemar/Cochran's Q.

^bWithin control group McNemar/Cochran's Q.

^cTwo-sample McNemar test for probability of change over time across the intervention and control groups (none observed).

^dPearson chi-squared/Fisher's exact test, examining the difference in proportions across the intervention and control groups at T1 (none observed).

^ePearson chi-squared/Fisher's exact test, examining the difference in proportions across the intervention and control groups at T3.

* $P < 0.05$.

** $P < 0.01$.

*** $P < 0.001$.

TABLE II. Twelve-Month Prevalence of Suicidal Behaviors by Intervention Status (Cross-Sectional Data at T1, T2, and T3)

Time point	Assignment	Suicidal behaviors															
		Any suicidal behaviors			Any suicidal ideations			Any suicidal plans			Any suicide attempts			Any non-suicidal self-injury			
		%	n	Past 12 months	%	n	Past 12 months	%	n	Past 12 months	%	n	Past 12 months	%	n	Past 12 months	%
Baseline (T1) (N = 9,980)	Intervention	14.2 ^{a***}	669	13.2 ^{a***}	619	5.8 ^{a**}	274	2.1 ^{c*}	101	3.0 ^{a*}	141	13.9 ^{a***}	652				
	Control	13.8 ^{b***}	726	12.7 ^{b***}	672	5.1 ^{b***}	267	1.6 ^{c*}	83	2.5 ^{b*}	131	12.7 ^{b***}	671				
Follow-up at 12 months (T3) (N = 13,698)	Intervention	10.5 ^{a***d***}	818	10.1 ^{a***d***}	785	4.8 ^{a***d***}	377	1.9 ^{d**q}	149	2.2 ^{a*}	175	10.5 ^{a***d***}	817				
	Control	8.5 ^{b***d***}	505	8.1 ^{b***d***}	481	3.5 ^{b***d***}	207	1.3 ^{d**}	76	1.9 ^{b*}	111	8.4 ^{b***d***}	495				
Follow-up at 5-6 months (T2) (N = 14,355)	Intervention	12.1	838	11.3	781	5.5	380	2.4	164	2.0	140	Note (e): proportions were not significantly different across groups					
	Control	11.5	852	10.7	796	4.8	360	2.2	165	1.9	140						

Abbreviation: SBQ-R, Suicide Behaviors Questionnaire—Revised.

^aZ-score test (overlapping samples)⁶, comparing within-group differences in proportions from T1 to T3—intervention.

^bZ-score test (overlapping samples)⁶, comparing within-group differences in proportions from T1 to T3—control.

^cPearson chi-squared/Fisher's exact two-sided test, comparing differences in proportions across groups at T1 (independent samples).

^dPearson chi-squared/Fisher's exact two-sided test, comparing differences in proportions across groups at T3 (independent samples).

^ePearson chi-squared/Fisher's exact two-sided test, comparing differences in proportions across groups at T2 (independent samples).

* $P < 0.05$.

** $P < 0.01$.

*** $P < 0.001$.

TABLE III. Suicidal Behaviors 6-Month Incidence and 12-Month Combined T2/T3 New Events (i.e., No Prior Lifetime Reported Events) by Intervention Status and Levels of R4 Use at T3 (linked data)

Assignment	Suicidal behaviors															
	Any suicidal behaviors (ASB)			Any suicidal ideations (ASI)			Any suicidal plans (ASP)			Any suicide attempts (ASA)			Any non-suicidal self-injury (ANSSI)			
	6-month incidence at T2 (no prior lifetime ASB)	Any new ASB past 12 months (from T1 to T3 no prior lifetime ASB)	%	6-month incidence at T2 (no prior lifetime ASI)	Any new ASB past 12 months (from T1 to T3 no prior lifetime ASI)	%	6-month incidence at T2 (no prior lifetime ASP)	Any new ASB past 12 months (from T1 to T3 No prior lifetime ASP)	%	6-month incidence at T2 (no prior lifetime ASA)	Any new ASB past 12 months (from T1 to T3 no prior lifetime ASA)	%	6-month incidence at T2 (no prior lifetime ANSSI)	Any new ASB past 12 months (from T1 to T3 no prior lifetime ANSSI)	%	
Intervention	5.6	77	6.8	128	5.6	79	3.2	50	3.8	82	0.8	13	1.1	17	1.4	30
Control	5.5	96	5.8	129	5.4	97	2.9	57	3.2	81	1.3	27	0.5	11	0.8	21
Intervention high use	3.9 ^{***}	28	5.3	35	4.1 ^{c*}	30	2.6	21	3.6	26	0.2 ^{b*}	2	0.5 ^{c*}	4	0.4 ^{c**}	3
Intervention intermediate use			7.1	41					7.1	42			0.7	5	1.7	12
Intervention low use	7.3 ^{c**}	44	8.2	46	7.4 ^{c*}	45	3.8	26	4.2	26	1.4 ^{c*}	10	1.6 ^{c*}	11	2.1 ^{c**}	13
Control	5.4	90	5.8	123	5.4	92	2.8	53	3.3	78	1.2 ^{b*}	25	0.5	9	0.8	19

^aOverall significance: Pearson chi-squared/Fisher's exact test.

^bHigh-use intervention group significant relative to control: Pearson chi-squared/Fisher's exact test.

^cHigh-use intervention group significant relative to low-use intervention group: Pearson chi-squared/Fisher's exact test.

* $P < 0.05$.

** $P < 0.01$.

*** $P < 0.001$.

difference in proportions across divisions at T1 or at T3 was statistically significant.

Cross-sectional within-division results also demonstrated smaller proportions of soldiers in both the intervention and control divisions who met the SBQ-R ≥ 7 threshold by T3 when compared to T1: net decreases of 3.4% ($P < 0.001$) and 4.3% ($P < 0.001$), respectively (Table II). The smaller proportion of soldiers in the control division cross-sectional sample who met the SBQ-R ≥ 7 threshold in comparison to the intervention division was significant at T3 ($P < 0.001$).

Suicidal ideations, plans, and attempts

Differences in proportions across divisions for ideations, plans, and attempts are reported both within and across divisions for linked data in Table I. Smaller proportions of soldiers in both intervention and control divisions reported suicidal ideation by T3 when compared to T1: net decreases of 2.1% ($P < 0.05$) and 3.3% ($P < 0.001$), respectively. Smaller proportions of soldiers in both intervention and control divisions reported suicidal plans, net decreases of 0.5% and 1.7%, respectively, but only the control division’s result reached significance ($P < 0.05$). The magnitude of the baseline to follow-up change across divisions for both ideations and plans was not statistically significant.

At T1, across the cross-sectional samples, a higher proportion of soldiers in the intervention group, in contrast to control, reported suicide attempts in the preceding 12 months ($P < 0.05$, Table II). Cross-sectional within-division results demonstrated smaller proportions of soldiers in both intervention and control divisions reporting suicidal ideations and suicidal plans by T3 when compared to T1: net decreases of 3.1% ($P < 0.001$) and 4.6% ($P < 0.001$) for ideations and of 1.0% ($P < 0.01$) and 1.6% ($P < 0.001$) for plans, respectively (Table II). The smaller proportion of soldiers in the control cross-sectional sample who reported suicidal ideations, suicidal plans, and suicide attempts in comparison to the intervention was significant at T3 (ideations/plans $P < 0.001$, attempts $P < 0.01$).

The 6-month incidence and proportions for new suicidal ideations, plans, and attempts during the 12-month study period are presented in Table III. Among findings from T2 linked data (pre-COVID), across division differences in the 6-month incidence of suicide attempts were not statistically significant, but subanalyses of T2 data demonstrated that higher-use R4 BNs reported significantly lower 6-month incidence of suicide attempts than the control division (Table IV, $P < 0.05$) or the lower-use R4 BNs ($P < 0.01$). The proportion of new suicide attempts during the 12-month study period was not statistically significant, neither across divisions nor by R4-use status.

Non-Suicidal Self-Injury

Based on analyses of T1-to-T3 linked data (Table I), smaller proportions of soldiers in both the intervention and control divisions reported NSSI by T3 when compared to T1: net

TABLE IV. Soldier-Reported Change in Suicidal Thoughts and Attempts post-COVID-19 by Intervention Status

	Linked data		Not applicable		Decreased		Stayed the same		Increased	
	Assignment	n	%	n	%	n	%	n	%	n
Thoughts about killing self since February 2020 compared to before COVID-19	Intervention (N = 1,251)*	1,125	89.5	1,125	2.6	33	4.6	47	2.9	36
	Control (N = 1,308)*	1,193	91.2	1,193	3.4	44	4.1	53	1.4	18
Attempts to kill myself since February 2020 compared to before COVID-19	Intervention (N = 1,213)	1,176	96.9	1,176	0.9	11	1.9	23	0.2	3
	Control (N = 1,265)	1,228	97.1	1,228	1.3	16	1.6	20	0.1	1
Thoughts about killing self since February 2020 compared to before COVID-19	Cross-sectional data	Not applicable			Decreased		Stayed the same		Increased	
	Intervention (N = 7,432)***	6,667	89.7	6,667	2.9	212	5.1	382	2.3	171
	Control (N = 5,682)***	5,217	91.8	5,217	2.8	157	4.0	229	1.4	79
	Intervention (N = 7,171)	6,902	96.2	6,902	1.4	99	1.9	137	0.5	33
Control (N = 5,486)	5,294	96.5	5,294	1.4	75	1.9	106	0.2	11	

Abbreviation: COVID-19, coronavirus disease 2019.

* $P < 0.05$.

*** $P < 0.01$.

*** $P < 0.001$.

decreases of 1.1% ($P < 0.05$) and 1.3% ($P < 0.01$), respectively. Neither the magnitude of this baseline-to-follow-up change across divisions nor the difference in proportions across divisions at T1 or T3 was statistically significant.

Cross-sectional within-division results also demonstrated smaller proportions of soldiers in both intervention and control divisions reporting NSSI by T3 when compared to T1: net decreases of 0.8% ($P < 0.05$) and 0.6% ($P < 0.05$), respectively (Table II). Differences in proportions across divisions for cross-sectional NSSI data are reported in Table II, and 6-month incidence and the proportion of new NSSI during the 12-month study period are presented in Table III.

COVID-19 and suicidal ideations and attempts

When asked about their experience with the frequency of suicidal thoughts since the COVID-19 pandemic (i.e., February 2020 thereafter), a statistically significant proportion of soldiers in the intervention group, in contrast to control, reported an overall increase in the frequency of suicidal thoughts during this time period, compared to before (Table IV). This was present in both linked ($P < 0.05$) and cross-sectional data ($P < 0.001$) (Table IV). Neither group reported appreciable change in suicide attempts since COVID-19.

DISCUSSION

The R4 intervention and control divisions both demonstrated significantly reduced total suicidal behaviors, total suicide risk (modified SBQ-R), suicidal ideations, and NSSIs in the 1-year period following R4 intervention as compared to the year before it. This pattern was clear for these outcome measures in both linked-longitudinal and cross-sectional samples. The magnitude of changes in these outcome measures among linked samples was not significantly different between divisions, indicating similar within-division effects for these outcomes in both locations. In intervention subgroup analyses, higher-use R4 BNs had a significantly lower incidence of total suicidal behaviors, suicidal ideations, and NSSI than lower-use R4 BNs during the pre-COVID (first 6 months) intervention period. This finding was only durable for NSSI by study completion.

Suicidal plan outcomes were mixed. The linked sample showed a statistically significant reduction in plans within the control division year to year, but the magnitude of the change in this outcome measure was not significantly different between the two divisions, indicating similar within-division effects for these outcomes in both locations. Both the intervention and control divisions demonstrated significant reductions within the cross-sectional samples. Subgroup analyses showed no significant differences in suicidal plans between high-/low-use BN groups and the control in either the first 6 months of the study or by study completion.

Explanations for the suicide plan findings include two likely contributors. First, the baseline suicide burden among the intervention division was higher than the control division. By T3, however, both divisions demonstrated comparable and statistically significant decreases in total suicide

risk (the SBQ-R measures lifetime occurrence of suicidal behaviors, frequency of suicidal behaviors in the previous year, likelihood of future events, etc.). In this context, a less steep decline in suicide plans within the intervention division may be explained by higher-to-lower severity behavior shifts and/or persistence of plan-related behaviors. The finding that there were no significant differences in plan incidence (new plans among those without preexisting suicidal plan behaviors) suggests those with preexisting suicidal plans accounted for a considerable portion of the plan differences across divisions.

Second, evidence suggests that COVID-19 more negatively impacted the intervention division than the control. Although it is unfortunate that the COVID survey addendum could not provide a more definitive determination for plans (only ideations and attempts were measured), the general uptrend in intervention division measures relative to COVID-related impacts could reasonably be expected to have had a similar contribution to plans as well.

Suicide attempt outcomes were mixed. Neither the intervention nor the control divisions demonstrated any significant within-division reductions in attempts year to year. Incidence findings in the pre-COVID period, however, were notable, and within the subgroup analyses, the higher-use R4 BNs had a significantly lower pre-COVID incidence of suicide attempts than both the control and lower-use R4 BNs. This finding was not durable by study completion.

Overall, both divisions demonstrated favorable reductions in total suicide burden during the course of the study. Both had relatively equivalent trends for total suicidal behaviors and total suicide risk (SBQ-R), and, among specific behaviors, both had relatively equivalent trends for suicidal ideations and NSSIs. Although both divisions demonstrated reductions in suicide plans, the control showed a more robust trend overall. Neither division demonstrated a significant reduction in suicide attempts, but subgroup analyses did show significantly less pre-COVID suicide attempt incidence among those with higher-use R4 relative to the control.

Despite significant reductions in multiple primary suicide outcomes within the intervention division, the similarity of these trends across multiple outcomes in both the intervention and nonequivalent control divisions obscures the establishment of a link between the R4 intervention and improvements in the outcomes of interest. Although R4 may have caused or contributed to positive results within the intervention division, the nonequivalent control division also performed very well without it.

Study Strengths and Limitations

Although multiple explanations are possible to explain the similarity of findings within and between the divisions, compensatory rivalry and instrumentation (longitudinal surveys of the population) in the setting of an un-blinded study are probable contributors.^{2,7,8} It is also possible that contamination of the control with portions of the R4 intervention occurred, as the R4 tools were published at the study's midpoint

(November 2019), but this is judged less likely given that major portions of the intervention (tool instructions) were lacking from that report.^{1,9} Although these are challenges from a methodologic perspective, a positive practical implication is that focused effort on the issue of suicide among mid-level Army leaders may be able to yield beneficial results through more than one mechanism.

It is noteworthy that the intervention division was able to demonstrate favorable reductions similar to the control because of two factors. First, although the divisions were highly comparable on many baseline demographic and outcome categories, the intervention division began the study with consistently and, in some cases, significantly higher suicide burden (suicide attempts) in its population. Second, the intervention division experienced significantly more increases in suicide burden (suicidal ideations) following COVID-19 onset than the control. Taken together, the intervention division was more significantly disadvantaged with regard to suicide burden (in total, lesser availability of leader time and resources per soldier), both at the beginning and end of the study, than the nonequivalent control. Potential contributors to these differences include slightly higher combat-arms representation in the intervention sample, disparate regional geographic pressures, and/or varying COVID-19 challenges.^{10–14}

Notwithstanding the first disadvantage, the midpoint (6-month) assessment allowed for some subgroup analyses free of the second confounder (COVID-19). Although the subgroup results do show some promise for the effectiveness of the R4 intervention in improving suicide outcomes of interest, they must be confirmed with additional studies. Self-selection bias—the possibility that BNs choosing to use R4 more often were already good at suicide prevention without R4 or had lower suicide burden from the outset—could not be ruled out in these 6-month analyses because of the 12-month pre-/post-study design.¹⁵ Additionally, the calculations are post hoc and subject to investigator biases.¹⁶ Future studies should be mindful to add a 6-month pre-assessment time point, take steps to blind investigators to subgroup assignments, ensure that the use variable developed herein is established as an a priori hypothesis, take steps to quantify attrition within units, and conduct any follow-on study during the post-COVID period.

CONCLUSIONS

There is no evidence of harm associated with the R4 intervention. The effectiveness of R4 as a function of R4 itself requires confirmatory study. In the absence of additional evidence or alternative interventions with a stronger evidence base, however, R4 is judged to be an improvement (no evidence of harm + weak evidence of effectiveness) over the status quo (no safety data or effectiveness studies) with regard to tool-based decision-making support for suicide prevention in the U.S. Army.

ACKNOWLEDGMENT

None declared.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Military Medicine* online.

FUNDING

This work is funded by the U.S. Army, Medical Technology Program Element—Warfighter Health Protection and Performance Standards (FY19).

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Curley JM, Penix EA, Srinivasan J, et al: Development of the U.S. Army's suicide prevention leadership tool: the behavioral health readiness and suicide risk reduction review (R4). *Mil Med* 2020; 185(5–6): e668–77.
2. Curley JM, Duffy FF, Kim PY, et al: Methodology of the U.S. Army's suicide prevention leadership tool study: the Behavioral Health Readiness and Suicide Risk Reduction Review (R4). *Mil Med* 2021; 186(3–4): 336–43.
3. Garvey Wilson AL, Hoge CW, McGurk D, et al: Application of a new method for linking anonymous survey data in a population of soldiers returning from Iraq. *Ann Epidemiol* 2010; 20(12): 931–8.
4. U.S. Congressional Budget Office: Department of the Army: the U.S. military's force structure: a primer. 1–27. Washington, DC: Congressional Budget Office; 2016; Available at <https://www.cbo.gov/sites/default/files/114th-congress-2015-2016/reports/51535-fsprimerbreakoutchapter2.pdf>; accessed January 14, 2022.
5. Osman A, Bagge CL, Gutierrez PM, et al: The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. *Assessment* 2001; 8(4): 443–54.
6. Derrick B, Dobson-Mckittrick A, Toher D, White P: Test statistics for comparing two proportions with partially overlapping samples. *J Appl Quant Methods* 2015; 10(3): 1–14.
7. Handley MA, Lyles CR, McCulloch C, et al: Selecting and improving quasi-experimental designs in effectiveness and implementation research. *Annu Rev Public Health* 2018; 39(1): 5–25.
8. Bärnighausen T, Tugwell P, Røttingen JA: Quasi-experimental study designs series—Paper 4: uses and value. *J Clin Epidemiol* 2017; 89(9): 21–9.
9. Ariel B, Sutherland A, Sherman LW: Preventing treatment spillover contamination in criminological field experiments: the case of body-worn police cameras. *J Exp Criminol* 2019; 15(4): 569–91.
10. Ursano RJ, Kessler RC, Naifeh JA, et al: Suicide attempts in U.S. Army combat arms, special forces and combat medics. *BMC Psychiatry* 2017; 17(1): 194.
11. Kessler RC, Stein MB, Bliese PD, et al: Occupational differences in US Army suicide rates. *Psychol Med* 2015; 45(15): 3293–304.
12. Curtin SC: State suicide rates among adolescents and young adults aged 10–24: United States, 2000–2018. *Natl Vital Stat Rep* 2020; 69(11): 1–10. Available at <https://www.cdc.gov/nchs/products/index.htm>; accessed January 14, 2022.
13. Ivey-Stephenson AZ, Crosby AE, Hoening JM, et al: Suicidal thoughts and behaviors among adults aged ≥18 Years - United States, 2015–2019. *MMWR Surveill Summ* 2022; 71(1): 1–19.
14. Pepper CM: Suicide in the mountain west region of the United States. *Crisis* 2017; 38(5): 344–50.
15. Humphreys K, Blodgett JC, Wagner TH: Estimating the efficacy of alcoholics anonymous without self-selection bias: an instrumental variables re-analysis of randomized clinical trials. *Alcohol Clin Exp Res* 2014; 38(11): 2688–94.
16. Srinivas TR, Ho B, Kang J, Kaplan B: Post hoc analyses: after the facts. *Transplantation* 2015; 99(1): 17–20.