Suicide interventions for American Indian and Alaska Native populations: A systematic review of outcomes

Tony V Pham, Anna Kawennison Fetter, Andrea Wiglesworth, LittleDove F. Rey, Micah L. Prairie Chicken, Michael Azaran, Amy Riegelman, Joseph P. Gone

ARTICLE INFO

Keywords:
American Indian populations
Suicide
Program development
Preventive interventions
Mental health disparities
Community-engaged research

ABSTRACT

Objective: A 2018 Center for Disease Control and Prevention report estimated that 22.1 per 100,000 American Indian/Alaska Native (AI/AN) individuals died by suicide, much higher than the overall U.S. rate of 14.2. To understand how to remedy this problem, we performed a systematic review in response to the following question: “What interventions work to prevent AI/AN suicide?”

Method: We adopted a broad inclusionary stance while searching, screening, and extracting data. Our search strategy yielded 1605 unique citations, and after screening 28 items met the set criteria.

Results: While participants from each study reported an improvement on at least one targeted measure, particularly along community-driven outcome measures, several methodological modifications arose to meet the ideals of both practice- and evidence-based research. For example, only 11 studies featured assessments that measured changes in direct suicide outcomes. Among these 11 studies, only four featured either a randomized or a non-randomized controlled trial. Furthermore, only one intervention produced consistent outcomes across several studies. Nevertheless, the results from our reviewed corpus were methodologically innovative and suggest an overall benefit to AI/AN communities.

Conclusions: The case for these interventions could be augmented through a variety of methodological advancements. Thus, we propose that future studies dismantle their interventions into underlying processes, evaluate these processes using direct, standardized measures of suicidal behavior, and incentivize AI/AN recruitment into research trials outside of Indian Country.

1. Introduction

A 2018 Centers for Disease Control and Prevention (CDC) report estimated that 22.1 per 100,000 American Indian/Alaska Native (AI/AN) individuals die by suicide, much higher than the overall U.S. rate of 14.2. Compared to other causes of death, suicide is the eighth leading cause among AI/AN individuals of all ages, and the second leading cause among AI/AN ages 10–24 years of age. Compared to White males, the rate of suicide does not increase into middle and older age but rather decreases (Centers for Disease Control, 1999–2018). These comparative differences in suicide circumstances beg the question of how effectively this unprecedented health crisis is being addressed. Before addressing this question, it is imperative to first consider the sociocultural roots of suicide within Indian Country.

During a long period of colonial dispossession, it was U.S. policy and practice to forcibly remove AI/AN individuals from their ancestral lands (Mohatt et al., 2014a; Sotero, 2006; Whitbeck et al., 2004). Once it became clear that AI/ANs had improbably survived dispossession and relocation, new policies were adopted to assimilate AI/AN individuals into the lower socioeconomic strata of U.S. society (Czyzewska, 2011;...
Sotero, 2006). As a part of this assimilation process, the U.S. prohibited cultural expression, and imposed Western cultural values, priorities, assumptions, and expectations at odds with AI/AN identity, history, and tradition (Wexler, 2009).

This resulted in a negative constellation of psychological and cultural sequela that has come to be known as “historical trauma” (Brave Heart & DeBruyn, 1998; Hartmann et al., 2019). Furthermore, a complex relationship has emerged between this historical trauma and risk/pro-\textit{tective} factors for suicide among AI/AN individuals (Wexler & Gone, 2012). For example, suicide risk factors for AI/ANs include alcohol and drug use, feelings of alienation, pressure to acculturate, discrimination, community violence, and exposure to the suicide of others. At the same time, protective factors include community control, cultural identification, spirituality, and family connectedness (Suicide Prevention Resour, 2013).

Despite growing awareness of these risk/pro-\textit{tective} factors, a severely under-funded mental health care system (Gone & Trimble, 2012) has yet to recommend and deliver interventions that specifically target AI/AN specific risk/pro-\textit{tective} factors (White & Kral, 2014). As a result, a disconnect between professional mental healthcare and AI/AN values has created disincentives for accessing community and mental health care services (Barlow & Walkup, 1998; Cunningham, 1993; Nelson et al., 1992; Novins et al., 1999), thus further perpetuating cultural conflict, health disparities, unemployment, lack of education, poverty, and geographical isolation (Doll et al., 2009).

In light of alarming rates of suicide, the scarcity of mental healthcare services, and the disconnect between mental healthcare needs and services available, multiple programs have emerged to resolve these issues. These interventions have been inspired by practice-based methodologies (e.g., flexibly constructed from the ground up to match the needs of specific AI/AN populations) and evidence-based interventions (e.g., selectively adopted from existing approaches that were developed and evaluated elsewhere, perhaps with cultural adaptations for new settings). Unfortunately, the methodological complexities to these practice- and evidence-based evaluations have often precluded formal meta-analytic summary and comparison.

Thus, the field of AI/AN suicide interventions requires at least some form of up-to-date analysis. Excluding scoping and narrative reviews, thus far, two studies have systematically reviewed suicide prevention programs for Indigenous peoples across Canada, Australia, New Zealand, and the US (the so-called CANZUS nations), including one among AI/AN youth (Harlow et al., 2014) and another across all age groups (Clifford et al., 2013). As of 2014, both described a heterogeneous corpus with study designs (e.g., non-randomized controlled trials) that did not lend themselves to estimations of causal efficacy. Based evaluations have often precluded formal meta-analytic summary and comparison.

Fortunately, given a recent surge in new, arguably more methodologically robust research over the past few years, the time is ripe for an updated systematic review that casts a broader net to inform development and implementation of future interventions and studies. To meet this demand, our systematic review sought to explore how interventions have addressed suicide among AI/AN populations using broad inclusion criteria that do not omit studies based on narrow demographics (e.g., focusing on AI/AN youth only) or methodological criteria such as time period (i.e., studies prior to 1981; Harvey et al., 1976), study design (i.e., case reports; Burt 1992; Gray & Muehlenkamp, 2010), and intervention type (i.e., discussion groups, service integration; Fleming 1994; Nebelkopf & Wright 2011).

2. Method

2.1. Transparency and openness

A multi-disciplinary, multi-institutional group of individuals comprised the research team. The team structured their preparation, execution, and report by following guidance outlined by Siddaway et al. (2019), the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Page et al., 2021), and the Cochrane Handbook for Systematic Reviews of Interventions (Cochrane, 2020). First, the team devised \textit{a priori} interests to systematize later choices in definitions, concepts, scope, and overall research design (e.g., the level of inclusionary flexibility). Second, the team composed their overarching research question, “\textit{What interventions work to prevent AI/AN suicide?”} All data and analysis are available upon reasonable request. This systematic review was not pre-registered. Further details regarding search strategy, article screening, and data extraction can be found below.

2.2. Search strategy

In June 2020, a social sciences librarian (AR) designed and deployed a primary search strategy in APA PsycINFO (Ovid). The complete primary search string is in Appendix A. To capture the cross-disciplinary nature of the proposed research question, she then translated the search string for eleven other bibliographic databases: Ovid Medline, EMBASE, CINAHL, ERIC via Ebsco, Bibliography of Native North Americans, Sociological Abstracts, Academic Search Premier, ProQuest Dissertations and Theses, PsyArXiv, SocArXiv, and SSRN. Furthermore, papers that outlined protocols for ongoing or future studies were followed up upon. This yielded one additional study published after the initial search (Tingey et al., 2020). Ultimately, the search strategy returned 1605 unique citations. These 1605 items were then exported into Rayyan, a web-based tool for managing systematic reviews (Ozzanzi et al., 2016). Rayyan afforded the systematic review team four immediate benefits. Rayyan could enable the team to (1) facilitate a clear record system, (2) generate a uniform but independent work environment for each collaborator, (3) mask and un-mask their decisions within this uniform work environment, and (4) help to ensure methodological rigor.

2.3. Article screening

Screening Criteria. Several inclusion/exclusion criteria guided the screening process through Rayyan. First, studies were included if they featured a sample that was at least 90% AI/AN or reported separate analysis for AI/AN individuals (including between group comparisons). Second, studies were included if they implemented an intervention (i.e., took deliberate action designed to bring about behavioral change) that was \textit{a priori} described as targeting suicide (mention of which would therefore be expected to appear in the early sections of the article). In other words, included interventions need not have resembled familiar suicide prevention efforts (e.g., reducing access to firearms) nor even have measured variables directly related to suicide. Rather, included studies must have introduced the intervention as intended to prevent suicide and measured variables at least indirectly related to suicide (e.g., hopelessness). This inclusion criterion allowed us to capture the broadest range of designed suicide interventions and reflects the epistemological design of AI/AN interventions that focus not just on immediate causal factors (e.g., substance use), but upstream prevention (e.g., purpose, belonging). Third, reported findings must result from or in association with the \textit{implementation} of the intervention (i.e., not literature reviews, systematic reviews, commentaries, process descriptions about previous or forthcoming studies, etc.). We did not exclude studies based on methodological rigor or specific sub-populations, interventions, controls, or outcome measures. Fourth, the article must have appeared in published, peer-reviewed journals.

Despite growing acceptance for including grey literature in systematic reviews (Golder, Loke, & Bland, 2010; Hartling et al., 2017; Trispodi, Barbui, & Cipriani, 2011), we focused on peer-reviewed sources to assure baseline quality in study reporting. Beyond this basic criterion, however, published studies were included regardless of underlying methodological rigor, as quantitative comparison was quickly ruled out owing to evident diversity across studies relative to stock indicators such as patient populations, interventions, controls, and outcomes (Fineout-Overholt &
Screening Process. Overall, screening occurred in two phases: (1) title/abstract screening and (2) full-text review for eligibility. Between July and August 2020, authors AW and LFR independently completed title/abstract screening of the 1605 unique items. Title/abstract screening yielded 684 items ($k = 0.83$). During September 2020, authors AnonymousTVP, AKF, and LFRAW independently completed full-text reviews for eligibility among these 684 items. Full-text screening yielded 28 items ($k = 0.80$). All disagreements were ultimately resolved by screener consensus. For example, one borderline case performed epidemiological analyses leading up to an intervention. This case was excluded on the grounds that its analyses were not used to evaluate the intervention.

Once all disagreements were resolved, the final corpus was reviewed to ensure that each included publication reflected the original inclusion and exclusion criteria without over- or under-inclusion. This final corpus was then exported into an open-source citation manager, Zotero. Each step in the searching and screening process was documented using the PRISMA (see Fig. 1).

2.4. Data extraction

To extract data from our final corpus, an extraction template was developed using guidelines by both Siddaway et al. (2019) and the Cochrane Handbook for Systematic Reviews of Interventions (Cochrane, 2020), with each extraction item reflecting a component of the underlying research priorities. The studies were extremely heterogeneous and so we refrained from conventional meta-analysis given the wide diversity in theoretical orientations, constructs, designs, methods, and outcomes (Baumeister & Prinstein, 2013; Siddaway et al., 2019). Moreover, scales or checklists were not utilized to evaluate scientific rigor because few of the studies adopted controlled designs that were necessary for assuring robust causal inference. Thus, formal evaluations of rigor would only have revealed that the vast majority of studies had “very low” or “low” quality of evidence.

3. Results

To structure our review, we highlight key features of the corpus as they related to our research question, “What interventions work to prevent AI/AN suicide?” To accomplish this, first, we deliver a broad overview of the corpus. Second, to afford insight into intervention outcomes, we attend to this literature using study design as an organizational framework.

Overall, our systematic review yielded a final corpus of 28 studies comprising 23 unique interventions. To evaluate these 23 interventions, researchers adopted a variety of study designs and outcomes. Eleven studies measured changes in suicide behaviors directly (i.e., suicide ideation, behaviors, attempts, deaths; see Table 1 for a complete breakdown of these eleven studies). Three of these 11 studies analyzed their outcomes using either a randomized or a non-randomized controlled trial, seven adopted a single group design (studies in which researchers assessed outcomes over time among a single group of participants who all received the intervention or some variant of the intervention) and one was presented as a case report.

In contrast to measuring changes in direct suicide behaviors, 26 studies measured suicide outcomes indirectly using proxy variables (e.g., alcohol abuse), of which nine measured both direct and indirect measures, and 17 measured indirect suicide variables only (see Table 2 for a complete breakdown of these 17 studies). One of these 17 studies analyzed outcomes using a non-randomized controlled trial, 14 adopted single study designs, and two were presented as case reports. Taken

![Fig. 1. PRISMA flow diagram.](image-url)
together, all studies reported improvement in at least one of their targeted outcomes (which was perhaps required for publication).

To further investigate these findings, we now review findings as classified by study design, while describing relevant methods and highlighting representative studies. Note that many authors reported non-statistically significant improvements in suicidal outcomes; although such results (by convention) are not considered findings, we include this information throughout our review (when reported by authors) so that it remains clear that these associations were in fact tested and reported.

### 3.1 Randomized controlled trial (one article)

Between May 2014 and June 2019, Tingey et al. (2020) conducted a RCT to evaluate the impact of a culture and entrepreneurial camp on youth behaviors. Participants randomized to intervention took part in ten lessons and six workshops about Apache culture and entrepreneurship whereas control participants took part in sports activities alone. The study randomized 394 middle and high schoolers (ages 13–16 years old) in a two-to-one ratio to intervention (n = 267) and control (n = 127) groups, respectively. As requested by community partners, they selected a two-to-one ratio to maximize benefit among all participants. The program assessed participants at baseline and over the course of a 24-month follow-up period (i.e., at the 6-, 12-, and 24-month mark) using the Youth Risk Behavior Survey (YRBS), an instrument validated for AI/AN youth behaviors. Participants randomized to intervention took part in ten lessons and six workshops about Apache culture and entrepreneurship.

Tingey and colleagues then evaluated the data using a Statistically significant increase; chi-square test statistics, cross-sectional, two-sample t-test; and mixed effects logistic regression. They found that participants who received the intervention reported a non-statistically significant reduction in suicide attempts but a statistically significant improvement in marijuana abuse (at all time points), fighting (specifically at the 24-month time mark), and school attendance (specifically at the six-month time mark).

To date, the study by Tingey and colleagues represents the only study that has implemented a RCT of an AI/AN suicide intervention. Interestingly, five non-RCT studies explained why randomizing to intervention or control was considered unfeasible or ill-advised by community representatives, who expressed discomfort around the impractical and inappropriate nature of randomization. For example, LaFromboise and Howard-Pitney (1999) noted that “a RCT was not feasible mainly due to [their] community partners’ preferences and the preliminary nature of the work” (p. 405). Furthermore, Allen et al. (2018) noted “that withholding a program was inconsistent with Yup’ik cultural values of inclusion; from a community perspective, if an intervention is thought beneficial, why would you randomly withhold it from some?” (p. 183). Similarly, Bartgis and Albright (2016) reported on how participants expressed reluctance to volunteer for a study that randomly assigned them to an experimental or control group.

### 3.2 Quasi-experiments (three articles)

Given the community hesitation surrounding randomization to treatment, other researchers proposed quasi-experimental designs as an alternative to gathering evidence about causal efficacy. Specifically, two research teams evaluated their interventions using a quasi-experimental design (see Tables 1 and 2). Of the two, LaFromboise and Howard-Pitney (LaFromboise & Howard-Pitney, 1994, 1995) assessed the impact of their intervention by measuring changes in direct suicide factors.
Between January to May 1990, LaFromboise and Howard-Pitney (LaFromboise & Howard-Pitney, 1994, 1995) recruited 106 participants for their 1994 study and 128 participants for their 1995 study. They evaluated the impact of the ZLSD program on several self-report measures (with some variation between studies): the Suicide Probability Scale (SPS), the Indian Adolescent Health Survey, the Beck Hopelessness Scale (BHS), the Symptom Check List-90-R, an adapted version of the Social Readjustment Rating Scale, and other custom scales designed to measure skill proficiencies as assessed by self-report or by judge/classmate report. Results were mixed across the two studies, particularly for overall suicide risk.

Per the 1994 study, LaFromboise and colleagues (LaFromboise & Howard-Pitney, 1994) did not collect pre-test measures from their control arm “due to teacher concerns regarding limited class time and the fears that discussion of suicide without instruction in the control classes would be harmful to students” (p. 113). Nevertheless, t-test analyses found that participants who received the intervention demonstrated a statistically significant pre-post improvement on the SPS and its sub-scale measures of suicidal ideation (SI), hopelessness, and hostility. Furthermore, a 2 (intervention, control) x 2 (pre-test, post-test) analysis of covariance on a modified version of the SI subscale found that, after controlling for pre-test differences, the intervention group had a statistically significant improvement when compared to the control arm. In contrast to the 1994 study, t-test analyses from the 1995 study (LaFromboise & Howard-Pitney, 1995) found a non-statistically significant improvement on the SPS but a significant improvement on the BHS and measures related to suicide intervention skills and problem solving, particularly during mild rather than serious suicide role play scenarios.

### Table 2
Studies that only measured changes in indirect suicide factors as organized by study design.

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Intervention Characteristics</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quasi-experiment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allen et al. (2018)</td>
<td>Qungavik Toolbox: Promote community ownership and individual, family, cultural protective factors through SUI prevention activities</td>
<td>Small effect ↑ in ‘reasons for life’ in a dose response relationship to attendance</td>
</tr>
<tr>
<td><strong>Single group designs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allen, Mohatt, Fok, Henry, and People Awakening Team (2009)</td>
<td>Qungavik Toolbox: Promote community ownership and individual, family, cultural protective factors through suicide prevention activities</td>
<td>↑ in youth reports of support and opportunities in their community in a dose response relationship to activity attendance</td>
</tr>
<tr>
<td>Barnett, Schmidt, Trainor and Wexler (2020)</td>
<td>Camp Pigaax: Connects youth to culture/mentors/Elders through activity and skill building camp</td>
<td>Men scored higher than women in perceived emotional negativity, interpersonal-, and self-worth</td>
</tr>
<tr>
<td>Bartgis and Albright (2016)</td>
<td>Kognito Gatekeeper Training: Train gatekeepers on identifying signs and symptoms of suicide through emotionally responsive online avatars</td>
<td>↑ in gatekeeper preparedness, likelihood, and self-efficacy</td>
</tr>
<tr>
<td>Cwik et al. (2014)</td>
<td>Celebrating Life: Provide support and referrals through a suicide surveillance system</td>
<td></td>
</tr>
<tr>
<td>Cwik et al. (2016c)</td>
<td>ASIST: Teach suicide first-aid skills through lectures, discussions, group simulations, and role-plays</td>
<td>↑ in knowledge and skill-based self-efficacy</td>
</tr>
<tr>
<td>Cwik et al. (2019)</td>
<td>Elders’ Resilience: Visit classrooms and help youth connect with their heritage, traditions, and culture through elder taught lessons</td>
<td></td>
</tr>
<tr>
<td>Doll and Brady (2013)</td>
<td>Sensory Tool: Sensory-based curriculum and activities that promote stress management for the purpose of suicide prevention</td>
<td></td>
</tr>
<tr>
<td>Fleming (1994)</td>
<td>MI Indian Studies Group: Discuss various MI issues through group discussions, culture/community activities, and formal presentations</td>
<td></td>
</tr>
<tr>
<td>Kerr et al. (2020)</td>
<td>Viewer Care Plan: Prepare adults for concerning social media; teach 3-step planning/response tool</td>
<td>5 patterns identified across 12 protective factors (&quot;Internal Orientation&quot; to self-efficacy and awareness of interconnection; &quot;External Orientation&quot; to giving, affection, praise, and family; &quot;Limits&quot; on alcohol abuse; &quot;Community/family&quot; and giving, affection, and praise; &quot;Low Protection&quot; from lack of exposure to all protective factors) with variations based on community/age</td>
</tr>
<tr>
<td>Mohatt et al. (2014a)</td>
<td>Qungavik Toolbox: Promote community ownership and individual, family, cultural protective factors through SUI prevention activities</td>
<td>↑ across several training efficacy measures (e.g., contacting youth, starting conversations, intervening, referring)</td>
</tr>
<tr>
<td>Muehlenkamp, Marrone, Gray, and Brown (2009)</td>
<td>Spiritual Advisory Committee: Connect students to tribal ceremonies based on their preference</td>
<td>YA community: small effect ↑ in peer protective factors and “reasons for life” in a dose response relationship to attendance</td>
</tr>
<tr>
<td>Wexler et al. (2017)</td>
<td>Youth Leaders Program: Teach a curriculum that addresses overall school climate</td>
<td>BT community: medium effect ↑ in individual and family factors, and “reflective processes” in a dose response relationship to attendance</td>
</tr>
<tr>
<td>Wexler et al. (2019)</td>
<td>PC CARIS: Facilitate learning circles that discuss the local/relevance/application of research</td>
<td>↑ in total knowledge scores</td>
</tr>
<tr>
<td><strong>Case report</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burt (1992)</td>
<td>Integrate self-perception, social context, and feminine experience through cultural art therapy</td>
<td></td>
</tr>
<tr>
<td>Gray and Muehlenkamp (2010)</td>
<td>Connect students with culture &amp; spirituality; reduce risk factors through support, gate keeper training, &amp; suicide prevention team</td>
<td></td>
</tr>
</tbody>
</table>

*Statistically significant unless otherwise specified.

Note. AI = American Indian; ET = Etlaum Tunnguur; ↑ = increase; MS = mental health; SA = suicide attempt; YA = Yupiit umaasalik navigation.
research. To compensate for the lack of a control arm, Allen and colleagues instead developed a variation of a stepped-wedge design, the dynamic wait-listed design (DWWLD), and compared outcomes between two communities by staggering their relative progress (i.e., dose level) throughout the intervention.

They enrolled 128 participants into the intervention and, instead of measuring for changes in direct suicide outcomes, they elected to measure variables that fit a previously tested multi-level theory of change model. These variables included (1) individual characteristics, (2) family characteristics, (3) community characteristics, (4) peer influences, (5) reflective processes, and (6) reasons for life (Allen et al., 2014). The researchers separated individuals across latent trait levels and performed hierarchical cluster analysis (mixed effects regression). Per these analyses, Allen et al. (2018) fitted a regression model that found upstream intervention effects on intermediate protective factors (i.e., individual, family, community, and peer influence factors) that led to down-stream protective factors (i.e., a statistically significant change in reasons for life but not reflective processes).

3.3. Single group studies (21 articles)

Owing to the many challenges that can present when working with rural, small, and culturally distinct populations, a majority of the reviewed studies adopted a single group design. For example, some researchers could not implement controlled or randomized study designs around the small, remote populations of circumpolar Alaska who base their day to day lives around seasonal changes and practices (Mohatt et al., 2014a). Other studies described financial barriers to undertaking experimental comparisons. For example, Barnett et al. (2020) noted how “a lack of funding and programmatic design challenges” prevented them from implementing an adequate control group to compare and account for other possible influences (p. 369). Ultimately, 21 studies spanning just over five decades (1967–2019) adhered to a single group design (see Tables 1 and 2).

3.3.1. Studies that measured changes in direct suicide factors

Six of these 21 studies included measurements of changes in direct suicide factors (see Tables 1 and 2). May et al. (2005) implemented Adolescent Suicide Prevention Project, a public health-oriented suicidal-behavior prevention team for youth (ages 10–24 years old), and assessed its impact between 1988 and 1992. This team was comprised of professional mental health staff and trained community lay providers. In light of access barriers and stigma related to mental health treatment, the team approached community members within more naturalistic community settings (e.g., outdoors, inside cars). Once connected, the team would provide psychoeducation, offer counseling, teach youth coping and adult parenting skills, advocate for the individual and relevant care services, and refer the individual to professional mental health services. While the number of recruited youth was unspecified, May and colleagues found that participants reported a 73 percent statistically significant drop in suicidal gestures and attempts, particularly among the younger age groups. In reflecting on the success of their program implementation, May and colleagues emphasized the importance of adequate staff development, vigilance, resource development, community relations, and robust administration.

Cwik et al. (2016a) recently published their outcomes from Celebrating Life, a community suicide surveillance system that referred at-risk individuals to professional health services and provided psychosocial support. Between 2007 and 2012, they recruited 2640 participants for the Celebrating Life program. Per their analyses, participants reported a non-statistically significant decrease in suicidal attempts and deaths relative to national rates. Furthermore, notable reflections arose following the completion of their study. First, their results suggested that building healthy relationships may prove more effective than restricting means to self-harm. Second, while dismantling different program effects proved challenging, Cwik et al. (2016a) emphasized the need for longitudinal outcomes and well-timed interventions.

Cwik et al. (2016b) also implemented New Hope, an intervention that followed up youth who presented to the emergency room because of a recent suicidal attempt. During follow-up, the New Hope team offered psychoeducation, skills, and assistance with treatment barriers. To evaluate New Hope, Cwik et al. (2016b) recruited 11 participants between 2009 and 2011. Their analyses found that participants reported a non-significant decrease on the Suicidal Ideation Questionnaire, but a statistically significant decrease on the Children’s Negative Cognitive Errors Scale and the Center for Epidemiological Studies-Depression Scale.

Harvey et al. (1976) evaluated a psychiatric consultation and social work services program at Mt. Edgecumbe School. Between 1968 and 1973 they surveyed 200 students who received the service and found that these students reported a non-statistically significant decrease in suicide attempts and a statistically significant decrease in expulsion and drop-out rates. Langdon et al. (2016) evaluated the Lumbee Rite of Passage (LROP) program, an intervention that sought to address SI and risk factors through a Lumbee cultural enrichment program. They recruited at least 38 participants who reported a non-significant decrease in SI among participants with at least two thirds attendance. In another study, Le and Gobert (2015) evaluated Restoring the Native American Spirit, an intervention that sought to deliver a mindfulness-based intervention. They recruited eight participants who reported a non-statistically significant decrease in SI, but a statistically significant increase in mindfulness, perceived skill acquisition, and social connections.

3.3.2. Studies that measured changes in indirect suicide factors

Nineteen single-group studies included measurements of changes in indirect suicide factors, with 15 of these 19 studies focused exclusively on indirect variables only (i.e., they did not also measure changes in direct suicidal behaviors, as reviewed in the earlier section). Specifically, 15 studies explored factors related to at-risk individuals, and four explored factors related to gatekeepers for at-risk youth (see Tables 1 and 2). Because these studies (with respect to their designs) are so limited in their ability to address efficacy questions, we illustrate each category with one study each.

A brief example of a study that focused on at-risk individuals includes an intervention developed for Alaska Native youth. Barnett et al. (2020) sought to enhance protective factors through Camp Pigaag, a 5-day culture camp in which youth received teaching and traditional storytelling from Elders, wellness practitioners, and guest presenters. These youth also participated in team-building and cultural group activities. While Barnett and colleagues did not measure for changes in direct suicide outcomes, they did assess the impact of Camp Pigaag through a multi-variate analysis of variance using the following pre-/post-intervention on various scales. They found significant improvements on measures related to affect, “belongingness,” and mastery of coping skills. Furthermore, they found that males had significantly higher scores than females on measures related to affect, self-perceived importance to others, and self-esteem.

A brief example of a study that focused on youth gatekeepers includes an assessment of gatekeeper training. Bartgis and Albright (2016) evaluated the impact of the Kognito Gatekeeper Training Simulation Program, a suicide prevention program that sought to train gatekeepers on how to recognize and intervene in potential suicidal behavior. Between 2011 and 2013, they recruited 86 match-paired participants (19 students, 41 faculty/staff, 10 high school educators, and 16 middle school educators) for their study. While they did not measure changes in direct suicide outcomes, they did collect pre-, immediately post-, and 3-month-post-intervention Gatekeeper Behavior Scale scores. Comparing 3-month-post to immediately post-intervention scores, gatekeepers reported a statistically significant increase in their perceived self-efficacy and likelihood to intervene but not in their preparedness. Comparing immediately post-to pre-intervention scores, however, revealed a significant increase in preparedness.
3.4  Case reports (three articles)

Three studies were case reports, each featuring girls seeking treatment at a health center (see Tables 1 and 2). Although case reports are thoroughly confounded with respect to causal inference, all three reported improvements in protective factors and overall suicidal risk. One case report by Kohrt et al. (2017) evaluated the impact of culturally adapted dialectical behavior therapy (DBT) on a 14-year-old Navajo girl hospitalized following a suicide attempt. At the end of her hospitalization, t-test analyses demonstrated significantly higher scores on the Reasons for Living Inventory for Adolescents. The treatment team then discharged the patient once her SPS scores decreased from a severe to moderate risk.

4. Discussion

We performed a systematic review of AI/AN suicide interventions to answer our research question, “What interventions work to prevent AI/AN suicide?” In light of previous systematic reviews that highlighted a lack of homogenous and methodologically rigorous data, we adopted a broad inclusionary stance while searching, screening, and extracting. To emphasize the most methodologically rigorous studies, we organized our results by study design (and reported all outcomes, including non-significant outcomes from assessments of direct suicide measures). This flexible approach yielded a total of 28 studies spanning from 1968 to 2019. Taken together, each study reported improvement on at least one of their targeted outcomes (likely necessary for publication), however only 11 studies included assessments for changes in direct suicide outcomes (with two measuring direct variables only). Among these 11 studies, only three ascertained their outcomes using either a controlled trial (one randomized and two non-randomized). One of these controlled trials reported statistically significant improvement on the SPS (LaFromboise & Howard-Pitney, 1994), another reported a non-statistically significant improvement on the SPS (LaFromboise & Howard-Pitney, 1995), and the third reported a non-statistically significant improvement on attempted suicides as measured by the YRBS (Tingey et al., 2020). All other studies either used non-controlled study designs or measured indirect variables only. Specifically, eight studies ascertained direct outcomes using either a single group or a case report study design, and 17 studies measured indirect variables only.

The reviewed studies suggest that researchers have crafted culturally sensitive, community-informed suicide interventions that have aimed to produce benefit for AI/AN populations. Although promising, methodological limitations rendered any general determination of which interventions actually “worked” elusive. Confidence in causal attributions between interventions and outcomes demands accurate estimates of treatment effects through RCTs, and if not RCTs then quasi-experiments that control for confounding factors. In this corpus, only four studies featured either a randomized or non-randomized controlled trial. Moreover, confidence in efficacy also depends on replication of findings across multiple studies that collectively demonstrate meaningful effect sizes. This process of replication is conventionally structured around a homogenous set of populations, interventions, controls, and outcome measures (i.e., the PICO framework).

In this corpus, only two interventions were featured in more than one outcome study: the ZLSD was evaluated in two articles and the Qungavik program was evaluated in four articles (see Tables 1 and 2). Of these two programs, only the Qungavik program produced consistent statistically significant outcomes, albeit for changes in indirect suicide outcomes in line with a multi-level theory of change model (i.e., related to reasons for life and social support). In view of these limitations, additional replications of rigorous, controlled outcome studies that demonstrate direct effects on AI/AN suicide would be required to definitively answer our research question. Despite pressing suicide statistics and multiple calls for action to ameliorate AI/AN suicide, alongside two prior systematic reviews and our current review of (now) 28 intervention studies from five decades of investigation, what could account for the continued absence of rigorous and replicated findings that could guide effective suicide prevention for these populations? The answer is multifold.

4.1  Challenges of AI/AN community outcomes research

AI/ANs have been small in number, dispersed around the nation, and deeply underrepresented in clinical research. Additionally, even for AI/AN populations, suicide is a low base rate phenomenon. Together, these realities harbor crucial methodological ramifications. First, many studies have focused their evaluation on short term outcomes and small sample sizes. This has led to widened confidence intervals, potential “treatment diffusion” across small communities (Dumville et al., 2006; Peckham et al., 2015), and an inability to capture an intervention’s long-term sustainability and impact. Second, the RCT, a study design that requires robust structuring in well-controlled research settings, is challenging to implement, particularly in small, remote, and resource-strapped communities. Recall that Allen et al. (2014) observed that rigid design elements were unsuited for remote circum-polar regions of Alaska where small populations were more fluidly responsive to seasonal changes. Finally, RCTs are limited in their ability to assess intervention efficacy outside of such strictly controlled conditions.

Beyond these ecological challenges, AI/AN communities can be understandably suspicious of researchers and research, whether because of a history of mistreatment by the greater health sector or because of prior research engagements that were irrelevant, disrespectful, or even exploitative (Glover et al., 2015; Gone et al., in press). In response, AI/AN communities have increasingly exercised their powers of sovereignty as Tribal Nations to regulate research. This has made AI/AN research more participatory and community driven than ever.

For example, within our reviewed corpus, 20 of the 28 reviewed studies featured some form of community input during development/implementation with six explicitly invoking community based participatory research (Israel et al., 2013).

While community-driven development and implementation promotes community empowerment and accountable research, some scholars have described authorities, gatekeepers, and advocates in Indian Country who have expressed reluctance to support so-called “gold standard” research approaches and designs (Allen et al., 2018; Bartgis & Albright, 2016; LaFromboise & Howard-Pitney, 1994; Mohatt et al., 2014a; Wexler et al., 2019). Instead, community input has occasioned the use of community specific measures and designs that are tailored to the AI/AN populations. These include adoption of various risk/protection models related to “transactional-ecological” approaches, family connections, community-mindedness, and a collectivist self-orientation (see the 17 included studies that measured intervention outcomes based on models such as work/school attendance/performance, social engagement, family functioning) (Allen et al., 2014; Beeker et al., 1998; Doll & Brady, 2013; Erickson et al., 1988; Goodkind et al., 2015; Hawe et al., 1997; Hodge et al., 2009; Merzel & D’Afflitti, 2003). Sometimes, this involved steering researchers away from RCTs to more equitably increase access to interventions by community members. Adoption of these alternative designs led to measurement of variables that largely did not overlap across different studies, making establishing meaningful comparison among them challenging. At the same time, these variables better aligned with the values and practices of the communities they aimed to benefit. More importantly, these variables, many of which measure functional impairment, offer scientific value because of their relative scarcity in the literature. Furthermore, multiple studies have weaved them into experimental designs that are both scientifically and culturally credible.

In the case of the culture and entrepreneurial camp by Tingey et al. (2020), for example, the community voiced apprehension at randomizing participants to a control. Thus, Tingey and colleagues alleviated the community’s concerns by flexibly introducing a two-to-one
randomization strategy that maximized participant exposure to positive benefit. Alternatively, other researchers chose to solve randomization concerns by conducting non-randomized controlled trials. Beyond controlled trials, other studies evaluated their outcomes using single group designs and case reports while still comparing the differential impact of their intervention across different communities. For employment of a dynamic wait-list design to compare two communities at different stages of the intervention, see Mohatt et al. (2014b). For exceptions outside of suicide prevention, see McDonell et al. (2021) or Venner et al. (2020).

4.1.1. Recommendations for AI/AN community outcomes research

Under these conditions, we propose the following recommendations to continue the evaluation of AI/AN suicide interventions. First, researchers should consider measuring suicidal behaviors directly. On the one hand, suicidal behaviors are by their nature rare events and tribal communities may object to their measurement. On the other hand, demonstrations of intervention impact on direct suicide variables can strengthen confidence in the causal relationship between an intervention’s activities and a reduction in suicide behavior. If researchers are concerned about the ethics of exploring suicidal outcomes, certain methodological designs may help to mitigate this concern. For example, upcoming work by O’Keefe et al. (2019) will accommodate community misgivings by centering their RCT around the Sequential, Multiple Assignment, Randomized Trial (SMART) design, an approach that uses multiple time points to randomize and expose all participants to intervention, and that adjusts assignments and study variables on the fly (Lei et al., 2012).

Second, researchers should consider adopting stock measures of suicide, even if indirect, so that some findings are comparable across different studies. For example, several well-known measures that have been validated for AI/AN communities include the Suicidal Ideation Questionnaire, the Hopelessness Depression Symptom Questionnaire, The Patient Health Questionnaire, and the Suicidal Behaviors Questionnaire. Third, from the perspective of research as a whole, investigators outside of Indian Country should be incentivized to recruit AI/AN participants into research trials more generally. For example, zero AI/AN individuals were represented across 342 RCTs for depression as reviewed by Polo et al. (2019). Similarly, in a comprehensive review of research on evidence-based mental health interventions for minoritized ethnorracial populations in the U.S., Miranda et al. (2005) could find no studies that included AI/ANs as study participants. Thus, concerns about applicability of RCT-studied interventions for AI/AN communities cannot be allayed without researchers’ careful evaluation of efficacy amongst the people who will later be served.

Finally, researchers should consider standardizing intervention development in terms of process and function rather than in terms of forms and mechanisms (Gone and Calf Looking, 2015). For example, the Qungavik intervention affords a toolkit of interventions that are codified by their basic functions so that these can be adopted piecemeal to promote comparison across separate studies. Subsequent review articles can then highlight these studies while improving on the methodological limitations inherent within this systematic review. For example, we adopted a broad and flexible approach to capture studies relevant to our research question. As our review yielded a heterogeneous body of studies, and we could not apply quality rating measures for a comparison of quality, or meta-analysis for comparisons of efficacy across similar studies. At the same time, we opted to focus on peer-reviewed research only and did not systematically search the grey literature. Furthermore, our approach to screening likely neglected studies that targeted a wide array of relevant (upstream) risk and protective factors for suicide because the authors did not explicitly link these to suicide. Lastly, we did not have space in this article to review the precise characteristics of each intervention (although this work is in preparation).

5. Conclusion

Death by suicide is remarkably high among AI/AN individuals within the U.S. Given this, the current systematic review hoped to understand, “What interventions work to prevent AI/AN suicide?” Results from the 28 articles revealed 23 distinct interventions with each study reporting an improvement on at least one targeted outcome. Nevertheless, questions about actual intervention efficacy remain open because of several methodological modifications. For example, only 11 studies included assessments for changes in direct suicide outcomes, and among these 11 studies, only three included outcomes derived from either a randomized or a non-randomized controlled trial. Furthermore, among the 23 reviewed interventions, only one produced consistent outcomes, albeit through the use of indirect measures of suicide. Many of these methodological modifications stemmed from the realities of collaborative research partnerships undertaken in AI/AN communities. Future outcomes research on suicide interventions in Indian Country will need to further identify innovative ways to develop study designs that both infer causality and resonate with community values and preferences.

Contributors

TVP, AKF, AW, LFR, AR, and JPG contributed to the overall conceptualization, data curation, formal analysis, and original draft preparation. All other authors reviewed, edited, and approved the final article.

Role of the funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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.

Acknowledgments

None.

Appendix A

Search Terms.

APA PsycINFO (Ovid).

1. Exp Self-Destructive Behavior/

2. Suicidal Ideation/

3. (suicide or suicides or suicidal or suicidality or "murder-suicide" or “homicide-suicide”).tw.

4. (“self-injur*” OR "self-harm*" OR automutilat* OR "self-destructive").tw.

5. ((self OR oneself OR myself OR themself* OR himself OR herself) adj2 (kill* OR harm* OR injur* OR hurt* OR mutilat*)).tw

6. OR/1-5

7. exp American Indians/

8. (“native american” OR "american indian" OR "alaska native" OR inuit*).tw.

9. alaska natives/ or inuit/

10. 7 OR 8 OR 9

11. 6 AND 10
Trevisi, C., Barbui, C., & Cipriani, A. (2011). Why it is important to include unpublished data in systematic reviews. Epidemiology and Psychiatric Sciences, 20, 133–135. https://doi.org/10.1017/S2045796011000217