Accounting for Diversity in Suicide Research: Sampling and Sample Reporting Practices in the United States

CHRISTINE B. CHA, PhD, KATHERINE M. TEZANOS, BA, OLIVIA M. PEROS, MA, MEI YI NG, MA, JESSICA D. RIBEIRO, PhD, MATTHEW K. NOCK, PhD, and JOSEPH C. FRANKLIN, PhD

Research on suicidal thoughts and behaviors (STB) has identified many risk factors, but whether these findings generalize to diverse populations remains unclear. We review longitudinal studies on STB risk factors over the past 50 years in the United States and evaluate the methodological practices of sampling and reporting sample characteristics. We found that articles frequently reported participant age and sex, less frequently reported participant race and ethnicity, and rarely reported participant veteran status or lesbian, gay, bisexual, and transgender status. Sample reporting practices modestly and inconsistently improved over time. Finally, articles predominantly featured White, non-Hispanic, young adult samples.

Suicide is one of the leading causes of death in the United States (Centers for Disease Control and Prevention [CDC] 2016a). Suicide thoughts and behaviors (STB) broadly include suicide death, suicide attempt, and suicide ideation. Approximately 1 million people in the United States attempt suicide each year, and more than 8 million experience suicide ideation each year (Crosby, Gfroerer, Han, Ortega, & Parks, 2011). STBs are not only life-threatening but also extraordinarily complex and challenging clinical outcomes of study.
To better understand STB, much research over the past 50 years has focused on the identification of psychological, psychiatric, demographic, and environmental risk factors of STB (Franklin et al., in press). Research on STB risk factors has tremendous potential to optimize efforts to identify who is most vulnerable and why (NAASP, 2014). A promising step in risk factor research has been the more frequent use of longitudinal designs relative to cross-sectional studies. The number of STB risk factor studies adopting a longitudinal design has increased over the past 50 years of research (Franklin et al., in press), and the average follow-up period across all of these STB risk factor articles now exceeds 9 years. Prospective studies now better adhere to the conventional definition of risk factor: a characteristic that is both associated with and temporally precedes the clinical outcome of interest (Kraemer et al., 1997). This marks significant improvement from studies featuring single time-point assessments, which could not identify time-varying, malleable, and potentially causal risk factors.

The exciting implications and improvements of STB risk factor research are tempered by modest findings and lingering methodological limitations (Franklin et al., in press). In the present investigation, we evaluate the state of this research and explore one of many potential areas for improvement: generality of findings. The question of which risk factor predicts STB must be accompanied by the information about exactly whom the risk factor pertains to. External validity, or generality, is central to this research because risk factors are not assumed to have a uniform impact across populations (Kraemer et al., 1997). A particular risk factor for women may not be a risk factor for men; for instance, non-impulsive aggression has been shown to be a risk factor for suicide death among women but not men (Dalca, McGirr, Renaud, & Turccki, 2013; Oquendo et al., 2007). Testing a single risk factor across different populations—without considering population-specific effects—may in part account for disparities in findings that have been observed throughout the field of psychiatry (Kraemer et al., 1997). Indeed, the concern around sampling and reporting of sample characteristics has long plagued multiple domains of psychology and behavioral sciences and applied fields (Gerfen, 1973; Henrich, Heine, & Norenzayan, 2010; Rosmarin, 2016; Sue, 1999). STB research may be no exception; the aforementioned meta-analysis of STB risk factors shows modest and heterogeneous effects across all the categories of risk factors (Franklin et al., in press). Despite methodological advances, this significant gap pertaining to generality remains.

A critical component to determining generality of findings is the nature of the study sample. Sample characteristics include sociodemographic factors such as age, sex, race, ethnicity, veteran status, and lesbian, gay, bisexual, and transgender (LGBT) status. We focus on two primary concerns around sample characteristics: whether reported samples represent vulnerable populations of interest, and whether sample characteristics are being reported in studies at all. Regarding the first concern, it remains unclear whether STB research samples are representative of the population as a whole or of those affected by STB. The concern around sample characteristics is especially poignant for research on STB, whose prevalence varies across demographic groups. For instance, minority status as defined by race (Bridge et al., 2015; CDC, 2016a), ethnicity (Kann et al., 2015; Peña, Matthieu, Zayas, Masyn, & Caine, 2012), veteran status (Kang et al., 2015), and/or LGBT status (Fergusson, Horwood, & Beautrais, 1999; Grossman & D’Augelli, 2007; Hatzenbuehler, 2011) is a notable marker of STB risk. These demographic differences do not seem to be otherwise accounted for by psychopathology such as depression or substance use disorders (O’Donnell, Meyer, & Schwartz, 2011), or environmental stressors such as bullying (Mueller, James, Abrutyn, & Levin, 2015). It would therefore be important that these demographic characteristics are represented in the sample to guide interpretation of
findings, regardless of whether they are a central focus of the investigation.

Regarding the latter concern, documentation of sample characteristics represents a key initial step toward interpreting and improving the generality of research findings. Despite the simplicity and importance of this information, fewer than 40% of applied psychology articles report race/ethnicity (Case & Smith, 2000). Without knowing at least the basic demographic makeup of a sample, the generality of a specific STB risk factor finding would remain unknown. As Beutler (1996, p. 898) states, “Although not the most important factor, it is reasonable to question whether psychologists will ever obtain a clear picture of the nature and psychological implications of demographic qualities in the absence of either reports of participant demographics or large numbers of systematic investigations of those distinctions.” Poor documentation is an easily addressable barrier, which when addressed would facilitate researchers’ awareness and appropriate evaluation of external validity.

The current investigation marks the first effort to empirically assess the state of diversity science within STB risk factor research. We seek answers to the following questions: (1) What proportion of these studies report sample characteristics? (2) Have practices in reporting sample characteristics changed over time? (3) What is the average demographic makeup of samples within this field of research? To address these questions, we conducted a systematic review of all STB risk factor studies from the past 50 years. Addressing these questions is aligned with recent calls to improve the state of diversity science in clinical psychology and psychiatry research (Rosmarin, 2016) and offer an opportunity to take stock of work thus far.

**METHOD**

**Literature Search**

We conducted a systematic search of longitudinal studies that examined prospective risk factors for STB through January 1, 2015, using PubMed, PsycINFO, and Google Scholar. This search yielded a total of 365 unique longitudinal studies published in print or online by January 1, 2015. The present review was conducted within a larger meta-analysis whose inclusion/exclusion criteria are specified in Franklin et al. (in press). Here, we drew results from 158 articles (43.29%) from the United States. Non-U.S. papers (56.71%) were excluded from analyses because of distinct cross-national practices of defining and labeling of race and ethnicity (Morning, 2006).

**Information Extracted**

Coders were trained to extract sample characteristics from each of the STB risk factor articles, using criteria adapted from Weisz et al. (in press). Specifically, coders extracted information from each article pertaining to the sample: average age, sex (% female), race (% White, Black, Asian/Pacific Islander, Indigenous/American Indian/Alaskan Native, Other/Multiethnic, Unknown), ethnicity (% Latino/Hispanic), veteran status, and LGBT status (% heterosexual, gay, lesbian, bisexual, transgender, or unspecified sexual minority). Weighted means were calculated for articles reporting these characteristics by group (e.g., suicide attempters vs. nonattempters). Of note, to optimize the number of articles accounted for, sample information was extracted based on

---

1This was confirmed by significant differences observed between the U.S. and non-U.S. papers in the frequency of reporting of race, $\chi^2 = 176.51, p < .001, \Phi = .72$, and ethnicity, $\chi^2 = 54.75, p < .001, \Phi = .40$, in this study. The U.S. and non-U.S. studies did not differ in their reporting of other sample characteristics of age, sex, and LGBT status, $\chi^2 = 0.03–3.10, ps = .08–.86, \Phi_s = .01–.10$.

2Ethnicity was defined in accordance with the U.S. Census, but the authors acknowledge other ethnic classifications (e.g., African American vs. Caribbean American) that impact suicide risk as well (Joe, Baser, Breeden, Neighbors, & Jackson, 2006).
individual characteristics (e.g., female/male, White/Black/Hispanic)—not based on overarching label (e.g., sex vs. gender; race vs. ethnicity). Excellent interrater reliability was achieved for categorical variables ($k_s = 0.90$–$1.00$) and for continuous variables (ICCs = $0.84$–$1.00$).

Data Analysis

We conducted descriptive statistics to assess the proportion of STB risk factor articles that report sample characteristics. Among the 158 U.S. papers, we calculated the frequency of reporting age, sex, race, ethnicity, veteran status, and LGBT status across articles. Finally, we conducted chi-square analyses to identify changes in reporting sample characteristics over four publishing time periods: Pre-1985, 1985–1994, 1995–2004, and 2005–2014. We then calculated the average demographic makeup of samples from those articles that did report sample characteristics and conducted ANOVAs to assess change over time.

RESULTS

Among all 158 STB risk factor articles, there were notable discrepancies in how frequently certain sample characteristics were reported (Figure 1). A large majority reported the sample characteristics of age (94.9%) and sex (88.6%). In contrast, fewer articles reported race (29.1%), veteran status (3.2%), and LGBT status (1.9%) were grossly underreported. Of note, the three articles reporting LGBT status were those that had specifically prioritized the identification of demographic risk factors, as opposed to environmental stressors or diagnostic or psychological risk factors.

Comparing articles over time, the rate of reporting sample characteristics appeared to modestly and inconsistently improve (Figure 2). The reporting of age has become increasingly common, $\chi^2 = 19.16, p < .001$, $\Phi = .35$, to the point where 100% of recent studies (2005–2014) now report sample age. Sex did not significantly increase over time, $\chi^2 = 3.03, p = .39, \Phi = .14$, but showed gradual improvements from 75.0% to 91.4% of articles reporting sex. The reporting of race and ethnicity has also become more common overall, $\chi^2_s = 13.63$–$23.45$, $ps = .001$–.003, $\Phi_s = .29$–.39, but were each driven by distinct era-specific changes. In the case of race, there was a dramatic increase reporting from before 1985 to 1985–1994, $\chi^2 = 6.04, p = .01, \Phi = .45$, which then plateaued and featured no significant increases across more recent eras, $\chi^2_s = 0.10$–$1.65$, $ps = .20$–.75, $\Phi_s = .16$–.03. In the case of ethnicity, there has been a significant increase each era starting from 1985–1994, $\chi^2_s = 4.94$–$7.39$, $ps = .01$–.03, $\Phi_s = .24$–.28. Of note, even in the most recent era less than half of the studies (43.0%) reported ethnicity. Articles reporting either veteran status or LGBT status were all featured in the most recent era of 2005–2014.

Finally, we examined articles reporting at least one of the aforementioned sample characteristics. Among these articles, the average sample makeup was relatively young ($M$ age = 30.0 years), slightly more likely to be female (53.6%), and much more likely to be White (70.2%) and non-Hispanic (89.6%) (Table 1). None of these characteristics significantly changed over time (e.g., Pre-1985 vs. 2005–2014), $F_s = 0.33$–$1.80$, $ps = 0.18$–.80. LGBT status and veteran

---

3The lack of consistency across these terms remains a valid point of contention (Beutler et al., 1996), but lies outside the scope of the present study.

4To gauge reliability, one master coder and five coders coded a random selection of 30 articles. Following the guidelines provided in Cicchetti (1994), interrater agreement ratings above 0.75 reflect "excellent" agreement.

5Estimates of racial minority representation may have been inflated due to the inclusion of some papers that specifically focus on a racial minority group (e.g., 100% Black, 100% Asian/Pacific Islander, and 100% American Indian/Alaskan Native).
Figure 1. Proportion of suicidal thoughts and behaviors (STB risk) factor articles reporting sample characteristics. Bars represent percentage of STB risk factor articles reporting specified sample characteristic. LGBT = Lesbian, gay, bisexual, and transgender.

Figure 2. Proportion of suicidal thoughts and behaviors (STB) risk factor articles reporting sample characteristics. Plots represent percentage of STB risk factor articles reporting specified sample characteristic within each era. LGBT = Lesbian, gay, bisexual, and transgender.
status were not averaged across studies nor compared across time since fewer than six articles had reported each of these demographic characteristics.

### DISCUSSION

We reviewed all the published longitudinal studies on STB risk factors over the past 50 years and identified three patterns of sampling and sample reporting practices. First, these articles grossly underreport demographic characteristics that mark STB risk such as race, ethnicity, veteran status, and LGBT status. Despite the tremendous amount of time and effort taken to uphold longitudinal study designs, the far simpler yet important task of reporting sample characteristics has yet to be standardized. The suicide literature is no exception, as fewer than half of the other clinical psychology studies report such demographic characteristics (e.g., Case & Smith, 2000; Mendoza, Williams, Chapman, & Powers, 2012), and yet are reported in less than 2% of longitudinal STB studies. Our findings confirm prior arguments made that nearly everything that is known about suicide risk and the LGBT population is based on cross-sectional survey or nonexperimental research (Haas et al., 2011).

Second, the practice of reporting sample characteristics has modestly and inconsistently improved over time. Now all articles report age, and nearly all report sex. The practice of reporting race, ethnicity, veteran status, and LGBT status is increasing, but this pattern is tempered by the fact that over 20% of studies still do not report race, over 50% of studies still do not report ethnicity, and over 95% of studies still do not report veteran or LGBT status. Importantly, this finding does not necessarily reflect the reporting patterns of the entire field of suicide research. There is burgeoning interest and greater attention paid to suicide risk as it pertains to these latter characteristics (e.g., Adrian, Miller, McCauley, & Vander Stoep, 2015; Dirkes, Hughes, Ramirez-Valles, Johnson, & Bostwick, in press). This finding is troubling not because there is no work being done in this area, but because the focus on these vulnerable

### TABLE 1

Average Sample Characteristics of Suicidal Thoughts and Behaviors (STB) Risk Factor Articles (N = 158)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Age (years)</td>
<td>30.0</td>
<td>14.4</td>
<td>39.4</td>
<td>9.4</td>
<td>30.0</td>
</tr>
<tr>
<td>Sex (% female)</td>
<td>53.6</td>
<td>22.3</td>
<td>51.3</td>
<td>23.2</td>
<td>52.4</td>
</tr>
<tr>
<td>Race (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>70.2</td>
<td>22.8</td>
<td>79.5</td>
<td>19.5</td>
<td>68.9</td>
</tr>
<tr>
<td>Black</td>
<td>20.4</td>
<td>21.4</td>
<td>22.6</td>
<td>21.2</td>
<td>24.1</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>8.6</td>
<td>22.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>American Indian/Alaskan Native</td>
<td>4.9</td>
<td>20.8</td>
<td>0.0</td>
<td>0.0</td>
<td>13.2</td>
</tr>
<tr>
<td>Ethnicity (% Latino/Hispanic)</td>
<td>10.4</td>
<td>7.2</td>
<td>12.2</td>
<td>10.1</td>
<td>10.2</td>
</tr>
</tbody>
</table>

Note. M, Mean; SD, standard deviation. None of these characteristics significantly changed across eras (e.g., Pre-1985 vs. 2005–2014), Fs = 0.33–1.80, ps = .18–.80. Veteran status and LGBT status were not included due to exceptionally few STB risk factor articles that reported these demographic characteristics.
groups has largely been driven by a small proportion of the field. The cross-sectional and emerging longitudinal work accounting for marginalized groups is critically important and could be easily spread to all methodologically rigorous work on STB done as they pertain to an array of risk factors. An initial and feasible step would be to implement sampling and reporting practices for race, ethnicity, and LGBT status as is already for age and sex.

Third and finally, articles reporting sample characteristics on average featured White, non-Hispanic young adults. This sampling pattern has remained relatively consistent across time and reflects much of the U.S. population toward the end of the last century. Indeed, the U.S. Census Bureau (2002) reports that the American population has been predominantly White (from 83.1% in 1980 to 75.1% in 2000), non-Hispanic (from 93.6% in 1980 to 87.5% in 2000), with a median age falling within the early adulthood (22.9 years in 1900 to 35.9 years in 2000). Despite the fact that earlier research studies may reflect the relative breakdowns of the population at the time of data collection, we remind researchers that the United States is projected to become significantly more ethnically and racially diverse in the coming years (e.g., by 2060 White non-Hispanic individuals will represent less than half of the population; U.S. Census Bureau, 2015) and that sampling practices will need to evolve accordingly to ensure research samples are demographically representative of the population (Kazdin, 1999). We also emphasize the importance of targeting known high-risk subpopulations. That is, although White and/or non-Hispanic young adults (30–35 years) feature some of the highest prevalence rates for suicidal thoughts and plans in their respective age group (CDC, 2016a; Crosby et al., 2011), they do not represent those who have been shown to have higher proportion of suicide attempt (e.g., American Indian and Alaskan Native populations within the same age range, White non-Hispanic adults older than 35 years; CDC, 2016a). This point parallels concerns voiced on an international scale, claiming that most suicide research is being conducted in regions and with populations that typically have lower suicide rates (Lopez-Castroman, Courtet, Baca-Garcia, & Oquendo, 2015).

It is important to note several limitations and caveats of the current review. First, we did not code several demographic characteristics such as socioeconomic status (SES) and religiosity, which have been linked with STB outcomes (Page et al., 2014; Rasic, Robinson, Bolton, Bienvenu, & Sareen, 2011). We have admittedly adopted a narrow definition of diversity and encourage future work to more comprehensively account for sample characteristics and explore key nuances within demographic groups that may more precisely distinguish STB risk (e.g., African American vs. Caribbean American; Joe et al., 2006). Second, there are many other factors that influence the generality of findings that were outside the scope of the present review (e.g., setting, specific measures). Finally, the present authors are not immune to the shortcomings observed within this review; we have omitted sample characteristics (e.g., LGBT status, SES) within our own published longitudinal studies (e.g., Cha, Augustein, Frost, Gallagher, D’Angelo, & Nock, 2016; Cha, Najmi et al., 2016; Cha, Najmi, Park, Finn, & Nock, 2010). We encourage more accountability across all the researchers, ourselves included, in efforts to elevate the level of research conducted throughout this field.

In sum, the study of STB risk factors has paid disproportionately little attention to sociodemographically diverse and marginalized populations. Efforts to address this oversight are increasing with time, beginning with the exceedingly simple step of reporting more than sample age and sex. Acknowledging who is being studied throughout the STB research will expand and add much needed nuance to our understanding of these clinical outcomes.
REFERENCES


KANN, L., MCMANUS, T., HARRIS, W. A., SHANKLIN, S. L., FLINT, K. H., HAWKINS, J.,


Manuscript Received: October 25, 2016
Revision Accepted: November 8, 2016